

STIC-Biotech/ChemLib

CRFE

145402

mg

From: Swope, Sheridan
Sent: Wednesday, February 16, 2005 8:59 PM
To: STIC-Biotech/ChemLib
Subject: 10/701,200

79280

For 10/701,200 pls search:

SID 6 against the NT and AA data bases.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

aa - 437

Beverly Shears

copy
this page

mg

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: 2/19/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence :# _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      February 24, 2005, 14:00:08 ; Search time 166 Seconds
              (without alignments)
              1018.158 Million cell updates/sec
```

Title: US-10-701-200-6

Sequence: 1 DVVTWPYHLTADIRFCHWF...FNSMLSEIGQPKGKVEVSH 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 21056922

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2284	100.0	437	5	ABB09168	Abb09168 Methylom
2	2284	100.0	437	5	AAE222299	AAe222299 Methylom
3	2284	100.0	437	5	ABG61553	Abg61553 High grow
4	1442	63.1	410	4	ABM45865	Abm45865 Propionib
5	1442	63.1	410	6	ABM42384	Abm42384 Propionib
6	659	28.9	184	8	ADS25315	AdS25315 Bacterial
7	306.5	13.4	350	8	ADN25455	Adn25455 Bacterial
8	282	12.3	488	5	ABB93840	Abb93840 Herbicid
9	281.5	12.3	485	3	AAG39480	Aag39480 Herbicid
10	270	11.8	448	6	ABU19327	Abu19327 Protein e
11	270	11.8	448	2	AAr71580	AAr71580 Flavaria
12	269.5	11.8	469	2	AAr71581	AAr71581 Oryza sat
13	263	11.5	426	2	AAr71583	AAr71583 Raphanus
14	262	11.5	485	2	AAr71579	AAr71579 Solanum t
15	261.5	11.4	496	7	ADCo7918	AdCo7918 Rice prot
16	261	11.4	345	8	ADBS30228	ADbs30228 Bacteria
17	257.5	11.3	500	5	ABBS3109	Abb93109 Herbicid
18	253	11.1	319	8	ADN20252	Adn20252 Bacteria
19	251	11.0	473	3	AAG36438	Aag36438 Arabidops
20	251	11.0	473	5	ABBS9145	Abb9145 Herbicida
21	251	11.0	537	3	AAG36437	Aag36437 Arabidops
22	250.5	11.0	357	8	ADN27016	Adn27016 Bacteria
23	247	10.8	366	3	AAG39481	Aag39481 Arabidops
24	246.5	10.8	987	6	ABRS2589	AbR52589 Protein e
25	246.5	10.8	987	7	ADK61892	AdK61892 Disease t

26	246.5	10.8	987	8	AD543568	Ad543568 Bacterial
27	243	10.6	462	5	ABB93203	Abb93203 Herbicida
28	240.5	10.5	320	4	AAU38192	Aau38192 Salmonell
29	240.5	10.5	320	6	ABU48101	Abu48101 Protein e
30	240.5	10.5	320	6	ABU47109	Abu47109 Protein e
31	240.5	10.5	320	6	ABU45064	Abu45064 Protein e
32	239.5	10.5	318	8	ADN26599	Adn26599 Bacterial
33	239	10.5	462	3	AA629869	Aag8969 Arabidops
34	239	10.5	489	3	AA629868	Aag8968 Arabidops
35	238.5	10.4	320	4	AAU34829	Aau34829 E. coli C
36	238.5	10.4	320	6	ABU28847	Abu28847 Protein e
37	238.5	10.4	320	8	AD645208	Ad645208 Bacterial
38	234.5	10.3	529	6	AAU50166	Abu50166 Protein e
39	234	10.2	307	6	AA647222	Aag47222 Arabidops
40	234	10.2	525	5	ABB93312	Abb93312 Herbicida
41	232	10.2	461	6	ABU48474	Abu48474 Protein e
42	232	10.2	462	4	AAU42927	Abu42927 Protein e
43	232	10.2	462	6	ABM39446	Abm39446 Propionib
44	229.5	10.0	295	8	ADN26502	Adn26502 Bacterial
45	229.5	10.0	320	6	ABU28456	Abu28456 Protein e

ALIGNMENTS

RESULT 1	
ABB09168	
ID	ABB09168 standard; protein; 437 AA.
XX	
AC	ABB09168;
XX	
DT	01-JUL-2002 (first entry)
XX	
DE	Methylomonas pyrophosphate dependent phosphofructokinase SEQ ID NO:16.
XX	
KW	Methylomonas; methanotropic; carbon metabolism; carbon flux pathway;
KW	trimesaldoxase; fructose biphosphate aldolase; KHG/KDPG aldolase; food;
KW	pyrophosphate dependent phosphofructokinase; phosphoglucomutase; feed;
KW	glucose 6 phosphate isomerase; 6-phosphogluconate dehydratase; methane;
KW	glucose 6 phosphate 1 dehydrogenase; enzyme; methanotroph; methanol;
KW	single carbon substrate; single cell protein; polysaccharide; thickener;
KW	isoprenoid; carotenoid pigment.
XX	
OS	Methylomonas sp.
XX	
PN	WO200220796-A2.
XX	
PD	14-MAR-2002.
XX	
PF	28-AUG-2001; 2001WO-US026730.
XX	
PR	01-SEP-2000; 2000US-0229906P.
XX	
PA	(DUPO) DU PONT DE NEMOURS & CO E. I.
XX	
PI	Koffas M, Odom JM, Norton KC, Ye RW;
XX	
DR	WPI: 2002-362250/39.
DR	N-PSDB; ABLS1513.
XX	
PT	New polynucleotide encoding a Methylomonas sp. carbon flux enzyme useful
PT	for altering carbon flow through methanotropic bacteria, utilized for
PT	production of single cell protein and commercially valuable
PT	polysaccharides.
XX	
PS	Claim 4; Page 66-68; 73pp; English.
XX	
CC	The present invention describes isolated polynucleotides (1) encoding a
CC	Methylomonas sp. carbon flux enzyme, consisting of: transaldolase;
CC	fructose biphosphate aldolase; KHG/KDPG aldolase; phosphoglucomutase;
CC	pyrophosphate dependent phosphofructokinase; glucose 6 phosphate
CC	isomerase; 6-phosphogluconate dehydratase; and glucose 6 phosphate 1
CC	dehydrogenase enzymes. (1) overexpression is useful for altering carbon
CC	flow.

flow through a methanotrophic bacteria, where the bacteria grows on a C1 carbon substrate of methane and methanol, and comprises a functional Embden-Meyerhof carbon pathway comprising a gene encoding a pyrophosphate dependent phosphofructokinase enzyme, where the bacteria is preferably Methylomonas 16a American Type Culture Collection (ATCC) PTA 2402, and where the carbon flux gene is: (i) over-expressed on a multicopy plasmid; (ii) operably linked to an inducible or regulated promoter; (iii) expressed in an antisense orientation; or (iv) disrupted by insertion of foreign DNA into the coding region. The manipulated methanotrophs are useful for the energy efficient conversion of single carbon substrates such as methane and methanol to commercially useful products in the food, feed and materials industries, and preferably for the production of single cell protein, and for producing polysaccharides, used as thickeners in food and non-food industries, and isoprenoid compounds and carotenoid pigments of various carbon lengths. The present sequence represents a Methylomonas pyrophosphate dependent phosphofructokinase from the present invention

Sequence 437 AA:

Query Match 100.0%; Score 2284; DB 5; Length 437;
 Best Local Similarity 100.0%; Pred. No. 6e-210;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVVTWPHYLTADIRFCHEWFLNPNFYTLNPKKVAITTAGGLAPCLNSAIGSLIERYTE 60
 DB 1 DVVTWPHYLTADIRFCHEWFLNPNFYTLNPKKVAITTAGGLAPCLNSAIGSLIERYTE 60
 QY 61 IDPSIEIICYRGYKGLLSDSYPTAEVRKKAAGVLOFGSGVIGNSVVKLTNVDCYKR 120
 DB 61 IDPSIEIICYRGYKGLLSDSYPTAEVRKKAAGVLOFGSGVIGNSVVKLTNVDCYKR 120
 QY 121 GLVKEGEDPQKVAADQVLDKGDVILHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTTV 180
 DB 121 GLVKEGEDPQKVAADQVLDKGDVILHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTTV 180
 QY 181 DNDVFPITKQSIGATTAAGCAEQGARYFNNAVANNANPRMLIVHVGWRNGWGLTAATAOERY 240
 DB 181 DNDVFPITKQSIGATTAAGCAEQGARYFNNAVANNANPRMLIVHVGWRNGWGLTAATAOERY 240
 QY 241 KLDRAEWLPELIGLTRYSEYHAHVFPDMAIDLEBAKRLREVMKVCVNIFFVSEGAGV 300
 DB 241 KLDRAEWLPELIGLTRYSEYHAHVFPDMAIDLEBAKRLREVMKVCVNIFFVSEGAGV 300
 QY 301 EAIYAEMQAKQEVPRDAFGHIKLDVAVNPKWFGSQFQOMIGAEXTLVOKSGYFAPASAS 360
 DB 301 EAIYAEMQAKQEVPRDAFGHIKLDVAVNPKWFGSQFQOMIGAEXTLVOKSGYFAPASAS 360
 QY 361 NVDDMRILIKSCADLAVECAFRESGVIGHDEBNGVLAIEFPRIKGGKPFNIDTDWPN 420
 DB 361 NVDDMRILIKSCADLAVECAFRESGVIGHDEBNGVLAIEFPRIKGGKPFNIDTDWPN 420
 QY 421 MLSEIGQPKGGKVEYSH 437
 DB 421 MLSEIGQPKGGKVEYSH 437

RESULT 2
 AAE22299 standard; protein; 437 AA.
 ID AAE22299 standard; protein; 437 AA.
 XX AAE22299;
 AC AAE22299;
 XX 07-AUG-2003 (revised)
 DT 25-JUL-2002 (first entry)
 XX
 DB Methylomonas 16a sp. phosphofructokinase pyrophosphate dependent enzyme.
 XX Carotenoid; isoprenyl pyrophosphate; antheraxanthin; aetaxanthin; diet;
 KW anti-oxidant; steroid; flavour; fragrance; electro-optic application;
 XX aquaculture; enzyme; phosphofructokinase pyrophosphate.
 OS Methylomonas sp.

XX W0200218617-A2.
 XX 07-MAR-2002.
 XX 04-SEP-2001; 2001WO-US027420.
 XX 01-SEP-2000; 2000US-0229858P.
 XX 01-SEP-2000; 2000US-0229907P.
 XX (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX Broscowicz PC, Cheng Q, Dicosimo DJ, Koffas M, Miller ES;
 PI Odom JM, Picataggio SK, Rouviere PB;
 DR N-PSDB; AAD35497.
 XX
 PT Producing carotenoid compounds e.g. antheraxanthin and aetaxanthin, by
 PT using microorganisms having a nucleic acid molecule encoding enzymes in
 PT the carotenoid biosynthetic pathway and which metabolize single carbon
 PT substrates.
 PS Claim 8; Page 104-106; 156pp; English.
 XX The invention relates to a method for producing carotenoid compounds. The
 CC method comprises a transformed metabolising host cell, comprising
 CC suitable levels of isopentenyl pyrophosphate and a nucleic acid molecule
 CC encoding an enzyme in the carotenoid biosynthetic pathway, under the
 CC control of regulatory sequences, and contacting the host cell with carbon
 CC substrate to produce a carotenoid compound. The method is useful for
 CC producing carotenoid compounds such as antheraxanthin and aetaxanthin, by
 CC using microorganisms having a nucleic acid molecule encoding enzymes in
 CC the carotenoid biosynthetic pathway and which metabolize single carbon
 CC substrates. The carotenoids have potent anti-oxidant properties useful in
 CC diet, and aquaculture elements. The carotenoids are also useful as
 CC intermediates in the synthesis of steroids, flavours and fragrances and
 CC compounds for potential electro-optic applications. The present sequence
 CC is Methylomonas 16a sp. phosphofructokinase pyrophosphate dependent
 CC enzyme used in the invention. (Updated on 07-AUG-2003 to correct OS
 CC field.)
 XX
 SQ Sequence 437 AA:

Query Match 100.0%; Score 2284; DB 5; Length 437;
 Best Local Similarity 100.0%; Pred. No. 6e-210;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVVTWPHYLTADIRFCHEWFLNPNFYTLNPKKVAITTAGGLAPCLNSAIGSLIERYTE 60
 DB 1 DVVTWPHYLTADIRFCHEWFLNPNFYTLNPKKVAITTAGGLAPCLNSAIGSLIERYTE 60
 QY 61 IDPSIEIICYRGYKGLLSDSYPTAEVRKKAAGVLOFGSGVIGNSVVKLTNVDCYKR 120
 DB 61 IDPSIEIICYRGYKGLLSDSYPTAEVRKKAAGVLOFGSGVIGNSVVKLTNVDCYKR 120
 QY 121 GLVKEGEDPQKVAADQVLDKGDVILHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTTV 180
 DB 121 GLVKEGEDPQKVAADQVLDKGDVILHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTTV 180
 QY 181 DNDVFPITKQSIGATTAAGCAEQGARYFNNAVANNANPRMLIVHVGWRNGWGLTAATAOERY 240
 DB 181 DNDVFPITKQSIGATTAAGCAEQGARYFNNAVANNANPRMLIVHVGWRNGWGLTAATAOERY 240
 QY 241 KLDRAEWLPELIGLTRYSEYHAHVFPDMAIDLEBAKRLREVMKVCVNIFFVSEGAGV 300
 DB 241 KLDRAEWLPELIGLTRYSEYHAHVFPDMAIDLEBAKRLREVMKVCVNIFFVSEGAGV 300
 QY 301 EAIYAEMQAKQEVPRDAFGHIKLDVAVNPKWFGSQFQOMIGAEXTLVOKSGYFAPASAS 360
 DB 301 EAIYAEMQAKQEVPRDAFGHIKLDVAVNPKWFGSQFQOMIGAEXTLVOKSGYFAPASAS 360
 QY 361 NVDDMRILIKSCADLAVECAFRESGVIGHDEBNGVLAIEFPRIKGGKPFNIDTDWPN 420

|||||
Db 361 NVDDMRLLKSCADLAVECAFRESGVIGHDEBNGNVLAIEFPRIKGGKPFNIDTWFNS 420
Qy 421 MLSEIGQPKGKVEVSH 437
Db 421 MLSEIGQPKGKVEVSH 437
RESULT 3
ABG61553
ID ABG61553 standard; protein; 437 AA.
XX ABG61553;
AC ABG61553;
XX 27-AUG-2002 (first entry)
XX High growth methanotrophic bacterial strain polypeptide #3.
XX High growth methanotrophic bacterial strain; C1 carbon substrate; enzyme;
XX High growth methanotrophic bacterial strain; C1 carbon substrate; enzyme;
XX methane; methanol; Embden-Meyerhof carbon flux pathway; 16S RNA;
XX pyrophosphate dependent phosphotransferase; nitrogen-containing compound;
XX ammonia; nitrate; nitrite; nitrogen; pigment; oxygen; landfill;
XX methane-containing environment; waste water treatment system; isoprenoid;
XX nitrous oxide; terpenoid; animal feed; carotenoid; exopolysaccharide.
XX Methylobionas sp.
XX WO200220728-A2.
XX 14-MAR-2002.
XX 28-AUG-2001; 2001WO-US026827.
XX 01-SEP-2000; 2000US-0229858P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX Koffas M, Odom JM, Schenzle A;
XX WPI: 2002-454358/48.
XX N-P8DB; ABR63332.
XX New high growth methanotrophic bacterial strain, useful for producing
XX single cell proteins, grows on a C1 carbon substrate, and comprises a
XX functional gene encoding in Embden-Meyerhof carbon pathway.
XX Claim 7, Page 85-87; 157pp; English.
XX The invention relates to a high growth methanotrophic bacterial strain,
XX which grows on a C1 carbon substrate e.g. methane and methanol, and
XX comprises a functional Embden-Meyerhof carbon flux pathway comprising a
XX gene coding a pyrophosphate dependent phosphotransferase enzyme or a 16S
XX RNA. The bacterial strain is useful for the production of single cell
XX protein and for the biotransformation of a nitrogen-containing compound,
XX e.g. ammonia, nitrate, nitrite or nitrogen. It is also useful for the
XX production of a feed product comprising a protein, carbohydrates and a
XX pigment and for reducing oxygen demand, for removing nitrates and
XX nitrates in methane-containing environments such as landfills, waste
XX water treatment systems or anywhere that methane, oxygen and nitrates are
XX present. The bacterial strain of the invention can be used as a
XX denitrifying agent for the conversion of nitrate or nitrite to nitrous
XX oxide with methane or methanol as a carbon source. It is also used in the
XX production of biomass including proteins, carbohydrates and a wide
XX variety of pigments (particularly for isoprenoid pigments for the purpose
XX of generating animal feeds), in production of terpenoid and carotenoid
XX compounds, useful as pigments and as monomers in polymeric materials and
XX in production of exopolysaccharides at high levels. Sequences ABG61551-
XX ABG61590 represent high growth methanotrophic bacterial strain proteins
XX of the invention
XX Sequence 437 AA;
Query Match 100.0%; Score 2284; DB 5; Length 437;

Best Local Similarity 100.0%; Pred. No. 6e-210;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DVVTWPHLTADIRFCWPFLENFYTLNNKRRKVAITLPAAGLAPCLNSAISLERYE 60
Db 1 DVVTWPHLTADIRFCWPFLENFYTLNNKRRKVAITLPAAGLAPCLNSAISLERYE 60
Qy 61 IDPSIEIICRYGGYKGLLDGSDYPVTAIEYRKKAGVLOPFGSVIGNSRYKLTNNVDCVR 120
Db 61 IDPSIEIICRYGGYKGLLDGSDYPVTAIEYRKKAGVLOPFGSVIGNSRYKLTNNVDCVR 120
Qy 121 GLVKEGDEPQKVAADQVYDIIHTTGGDDTNTAADLAAFLARNNGYGLTVIGLPTV 180
Db 121 GLVKEGDEPQKVAADQVYDIIHTTGGDDTNTAADLAAFLARNNGYGLTVIGLPTV 180
Qy 181 DNDVFPFKOSLGAMTAEGGARYFNNVVAENNAENPMLIVHEVMGRNGMLTAATQOER 240
Db 181 DNDVFPFKOSLGAMTAEGGARYFNNVVAENNAENPMLIVHEVMGRNGMLTAATQOER 240
Qy 241 KLIDRAEWLPELGLTRESYEVAVPEMAIDLEAEAKRLREVMKVCNIFVSEGAV 300
Db 241 KLIDRAEWLPELGLTRESYEVAVPEMAIDLEAEAKRLREVMKVCNIFVSEGAV 300
Qy 301 EAIIVAMQAKGOEVRPDAFGHIKLDVNPCKKFGGQFQOMIAEKTLYOKSGYFAPASAS 360
Db 301 EAIIVAMQAKGOEVRPDAFGHIKLDVNPCKKFGGQFQOMIAEKTLYOKSGYFAPASAS 360
Qy 361 NVDDMRLLKSCADLAVECAFRESGVIGHDEBNGNVLAIEFPRIKGGKPFNIDTWFNS 420
Db 361 NVDDMRLLKSCADLAVECAFRESGVIGHDEBNGNVLAIEFPRIKGGKPFNIDTWFNS 420
Qy 421 MLSEIGQPKGKVEVSH 437
Db 421 MLSEIGQPKGKVEVSH 437
RESULT 4
AAU45865
ID AAU45865 standard; protein; 410 AA.
XX AAU45865;
XX 27-FEB-2002 (first entry)
XX Propionibacterium acnes immunogenic protein #6761.
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX WO200181581-A2.
XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-US012865.
XX 21-APR-2000; 2000US-0199047P.
XX 02-JUN-2000; 2000US-0208841P.
XX 07-JUL-2000; 2000US-0216747P.
XX (CORI-) CORIXA CORP.
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI: 2001-616774/71.
XX N-P8DB; AAS59528.
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris.

PS Example 1; SEQ ID NO 7060; 1069pp; English.

CC Sequence AANU9105-AU66017 represent Propionibacterium acnes immunogeni
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and osteomyelitis), warts, joints and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: the sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 410 AA;

Query Match	63.1%	Score 1442	DB 4	Length 410
Best Local Similarity	68.6%	Pred. No. 3.5e-129		
Matches 271; Conservative	54	Mismatches	0	Gaps 0

Qy	33	KKVALITITAGGLAPCNSAIGSLIBERTIEDIPSEITCYAGGYGGLLGSYPPTAEVRKK	92
Db	15	KKVALLITGGFPACSTALISGLIÖRTTEVAPEVEIIAYHKGIEGLKDFLEVTDTVRKN	74
- Qy	93	AGVLÖFGSGVGNISRVKLTNYKDCYKRGJLYEGEDBOPKVAADOLVKDDGVDILHTIGDD	152
Db	75	AEILKRFPGSSPIGNISRVKLTNAADLVKRGJVAEGDDPLKVAADRLVADGVDILHTIGDD	134
Qy	153	TNTTAADILAAFLARNNYGLTVIGLPEKTVNDVYFPIKÖSIGAMTAEOGARFYFNVAENN	212
Db	135	TNTTAADILAAFLAENNGLTVVGLPXTINDVVPPIQSIGAMTAEOGSRFAQNIIGEHN	194
Qy	213	ANPRLLIYHEVWGRNCGMTLTAATOERYKLLDBAEMLPELGITRSEYEVYNAVFPBEMALD	272
Db	195	SGSRMLIYHEVWGRNCGMTLTAATAYREMLTLQOMLPEIGLSKKMDYHAAVVPBEAHID	254
Qy	273	LEAEAKRLREVDKDYDCVNI FVSEGGAEVAIYAEOMAKOEVRDPAFGHILKDVAWPKX	332
Db	255	LEAEAKRLREVDKDYDCVNI FVSEGGAEVAIYAEOMAKOEVRDPAFGHILKDVAWPKX	314
Qy	333	FGEOPQOMIGAECTLYVOKSGYFAPARASNVDDMRILIKSCADILAVECAFRESGVIHDED	392
Db	315	FGKQPADKLGAOKXWVQKSQGYFSRASAENSEADELIGRCTDLAVDCAIAGKTVIGODEB	374
Qy	393	NGNVLRAIEFPRIKGGKPFNIDITDWMNSMLSEIQ	427
Db	375	NGDILTINIAFDRIKGGKPEDTQOPFTALSEIQ	409

RESULT 5
ABM42384
ID ABM42384 standard; protein; 410 AA

AC ABM42384 ;

DT 20-OCT-2003 (first entry)

DE *Propionibacterium acnes* predicted ORF-encoded polypeptide #7060.

KW Acne vulgaris; antibeborrhoeic; dermatological; antibacterial;

KW immunostimulant; immune response; vaccine.

XX

OS Propionibacterium acne.

PN WO2003033515-A1.

PD 24-APR-2003

PF 11-OCT-2002; 2002WO-US032727

PR 15-OCT-2001; 2001US-00978825

PA (CORI-) CORIXA CORP.
XX

PI	Mitcham JL,	Skeiky YAW,	Persing DH,	Bhatia A,	Maisonneuve JL;
PI	Zhang Y	Wang S	Tan S	Idrees MT	Benson DR
PI					Jones P
PI					Carter D

Pl Barch B, Vallieve-Douglabs J;
XX

DR WPL; 2003-381/89/36.
DR N-PSDB: ACE64457.

AA
PT
New Propionibacteriu

PT or for stimulation

PS Example 1; SEQ ID NO 7060; 1481pp; English

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

CC additionally encompasses expression vectors and host cells comprising a

invention: a method for stimulating an immune response specific for a P. acnes polypeptide and an isolated T cell population comprising T cells prepared via this method; a vaccine composition (comprising P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of P. acnes in a patient; and a method for inhibiting the development of P. acnes in a patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridisation. The vaccine composition is useful for the stimulation of an immune response against P. acnes, or for treating acne, and the kit is useful for performing a diagnostic assay. The present sequence represents a polypeptide predicted to be encoded by an ORF (open reading frame) contained within the P. acnes polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WPO at <http://wipo.int/pub/> published pct. sequences

SQ Sequence 410 AA;

Query Match	63.1%	Score 1442	DB 6	Length 410	
Best Local	68.6%	Pred. No. 3.5e-129			
Matches 271	Conservative 54	Mismatches 70	Indels 0	Gaps 0	

QY	33	KKVALLTGGALPCLNNSAIGSLIEFTTEIDPSIEIILCYGQYKGLLGSGSPYTAELVRKK	92
QY	15	KKVALLTGGFARCLSTASISGLIQRYTTEVAPEVEIIAYHGVGEGLLKGDFLEATDTYRKN	74
Db	93	AGVLTREFGSGYI GNSRVKLTNYKDCYKRLVKEGEDPQCVAAADQLVKQGYDILHTIGDD	15
QY	75	AEILRFGGSP GNSRVKLTNAAADLVKGLVAEGDDPLCYAADRLVADGVDILHTIGDD	13
Db	153	TNTAAADLAALFARNNGYLTVIGLPRKYNDVDFPIKQSLGAMTAEOGARFPNNVAENN	212
QY	135	TNTTAADTLAAYLAENNGYLTIVGLKRTIDNDVVPTRQSLGAMTAEOGSRFAQNIIVEHN	19
Db	213	ANPRMLIYHEVWGRNCGMLTAATAOEYRKLDPRAEMLPELGITRSEYEVHAFVPEMAID	272

Db 195 SGSRMLVHEWGRNCGMLTAATPAKYREMLDTQQLPEIGLSKKAWDVHAYVPEAHID 254
 QY 273 LEAEAKRLREVMKDVCNIFVSEAGVEAIVAEOMQKQGEVPRDAFGHIKLDVANGKW 332
 Db 255 LEAEARLKNVMDVGVNITFLSEAGLDLIEEMEDQGEVRDPEGHVKKLVNPGAW 314
 QY 333 FGEQFQOMTGAETLVOKSGYFAPASAVNDMLIKSCADLAVECAFRRSEGVIGHDED 392
 Db 315 FGKQFADKLGAEKVMYQKSGYFGRSAASNEADLELIGRCTDLAVDCLAGKTGVIGQDER 374
 QY 393 NGNVLAIERPRIRKGGKPFNIDTDWFSMLSEIQ 427
 Db 375 NGDTLTNIAPDRIRKGGKPFDTQPFWTAMLSEIQ 409
 RESULT 6
 ADS25315
 ID ADS25315 standard; protein; 184 AA.
 XX
 AC ADS25315;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #14348.
 XX
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI, 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 14348; 122pp; English.
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 184 AA;
 XX
 Query Match 28.9%; Score 659; DB 8; Length 184;
 Best Local Similarity 67.9%; Pred. No. 1.4e-54;
 Matches 125; Conservative 27; Mismatches 32; Indels 0; Gaps 0;
 QY 33 KKVAILTAGGLAPCLNSAIGSLIERTEIDPSIEITCYRGKGLLGSYPYTAVERKK 92
 Db 1 QKVAMLTAGGLAPCLSSAVGGLIERYSIDIAPEIDIAYRSYGCVLLGERIBITKDMREK 60
 QY 93 AGVLGRFGSVIGNSRVKLTNNVDCVKGIVKGEDEPQKVAADQIVKDGVDLHTTGGDD 152
 Db 61 AHULHRYGSGPIGNSRVKLTNAADCAKRGIVKGEDEPQKVAADQITTIHTTGGDD 120
 QY 153 TMTAADLAAFLARNYGLTVIGLPKTVNDVFPPIKQSIGMTAAEQGARFPMNVVAENN 212
 Db 121 TMTAADLAAVIGANGYDLTVGLPKTVNDVFPPIKQSIGMTAAVGAFAFPMNVNEQS 180
 QY 213 ANPR 216
 Db 181 AAPK 184
 RESULT 7
 ADN25455
 ID ADN25455 standard; protein; 350 AA.
 XX
 AC ADN25455;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #8108.
 XX
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI, 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 8108; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 350 AA;

Query Match 13.4%; Score 306.5; DB 8; Length 350;

Best Local Similarity 26.5%; Pred. No. 2.6e-20;

Matches 104; Conservative 63; Mismatches 145; Indels 81; Gaps 14;

QY 23 FNPYTLNKKKVAITLTAAGLAPCLNSAIGSLIERTEIDPSIEIICVGGYKGLLGDPS 82
 DB 4 FKEYTOM---KVGVLITGGDCGGLNAVIRAVARRKIK-EYGEYEFVFGDMRGPLEGDT 58
 QY 83 YEVTAVERKKAGVLOFGGSGVIGNSRVKLTNVKDCVKGGLVKEGEDPQKAAADQLVKQGV 142
 DB 59 MFLDIEAVR-GILPR-GGTLIGSSRTNLMKIEGGER-----VQDNNAALGV 103
 QY 143 DILHTITGGDDTAAADLAFLARNNYGLTVIGLPTVDNDVFPPIKQSIGAWTAAGQAR 202
 DB 104 DALVAIGGEDTLGVARQL-----HDHGVNVVGVPEKTINDLNATDYTFGFDVAIVATE 157
 QY 203 YEMNVVAENNANPRMLIVHEVMGNRCGWTL-----AATQOYRKLLDRAEMLPELGITRE 257
 DB 158 AIDRLHTTAESHHRALVV-EVMGRHAGMTALHAGMAAGA----- 195
 QY 258 SYEVHAVFVPEMAIDLEAEAKRLREVMDCVNIIFVSEAGAVEAIVAEQAQGEVPRD 317
 DB 196 ----NVLILPERPFDIDEVAVAIIESFRKTNVAPIIVAAAGAHK-----EGQLTLASAEKD 247
 QY 318 AFGHTKLDVAVNPGKMFGEQFAQMIGAETLVOKS---GYFAPASASNVDMMLIKSCADL 374
 DB 248 SFGHVRILGSI-----GQRLAEIEIARTGKEARSVVLGHVORGTPSAPFDVLAIRLGIH 301
 QY 375 AVECAFRRESGVIGHDENGNV--LRAIFPRI 405
 DB 302 AITAV-----HDKDFGKMAVNLKGTIVAV 325

RESULT 8

ABB93840

ID ABB93840 standard; protein; 488 AA.

XX ABB93840;

DT 31-MAY-2002 (first entry)

XX

DE Herbicidally active polypeptide SEQ ID NO 3051.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP009892.

XX 28-AUG-2001; 2001WO-EP009892.

XX (FARB) BAYER AG.

XX Tiejfen K, Weidler M;

XX WPI: 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 CC organisms.

PS Claim 5; SEQ ID NO 3051; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides

XX Sequence 488 AA;

Query Match 12.3%; Score 282; DB 5; Length 488;

Best Local Similarity 26.8%; Pred. No. 9.5e-18;

Matches 112; Conservative 60; Mismatches 158; Indels 88; Gaps 18;

QY 37 ILTAGLAPCLNSAIGSL-----IERTEIDPSIEIICVGGYKGLLGDSPYPTA 87
 DB 96 IYTCGLCPGNTVIREVSSLSYWGVRILGIDVSL-----GGYGFYAKNTIPLNS 149
 QY 88 EVRKKAAGVLOFGGSGVIGNSRVKLTNVKDCVKGGLVKEGEDPQKAAADQLVKQGV 147
 DB 150 KV---VNDIHKRGGTLIGTSR-----GSHDTNKL-VDSIQRGINQYI 189
 QY 148 IGGDDTAAADLAFLARNNYGLTVIGLPTVDNDVFPPIKQSIGAWTAAGQARVF--M 205
 DB 190 IGGDGTGAGASVIFPEIRRRRLKVAVGIPTKIDNDIVIDKSPGFDPAVEAQAQAINAA 249
 QY 206 NVVAENNANPRMLIVHEVMGNRCGWTLTAATQOYRKLLDRAEMLPELGITRESYVAHV 265
 DB 250 HVEAASNSNGIGFV--KLMGRYSGIIA-----MYATLASRVDCC 288
 QY 266 VPEMAIDLEAE-----KRLREVMDCVNIIFVSEAGAVEAIVAEQAQGEVPRD 318
 DB 289 IPESPFYIEGEGGLFEPFERLKHGHNV-----IYLAGAAGDLMCKMES-----TPMDA 340
 QY 319 FGHITKLDVAVNPGKMFGEQFAQMIGAETLVQ---KSGYFAPASASNVDMMLIKSCADL 374
 DB 341 GSKKLKQV--GLWLSQSIKHOFKKNKVMNLKIDPTMYIRAVPSNADSNVCTLLAQ 398
 QY 375 AVECAFRRESGVIGHDE--NGNVIRALIEFPRIKSGKFPNIDTD--WFSMILSEIQ 428
 DB 399 AVHGAM--AGYTGYSGLVNGR-QTYIPFYRIETQNNVITDRMMAR-LISSTNQP 451

RESULT 9

AAG39480
ID AAG39480 standard; protein; 485 AA.
XX
AC AAG39480;
XX
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 48854.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140655P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142054P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145813P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147433P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.

```

PR 23-AUG-1999; 99US-0149902P
PR 23-AUG-1999; 99US-0149930P
PR 25-AUG-1999; 99US-0150566P
PR 26-AUG-1999; 99US-0150884P
PR 27-AUG-1999; 99US-0151065P
PR 27-AUG-1999; 99US-0151066P
PR 27-AUG-1999; 99US-0151080P
PR 30-AUG-1999; 99US-0151303P
PR 31-AUG-1999; 99US-0151388P
PR 01-SEP-1999; 99US-0151930P
PR 07-SEP-1999; 99US-0152363P
PR 10-SEP-1999; 99US-0153070P
PR 13-SEP-1999; 99US-0153758P
PR 15-SEP-1999; 99US-0154018P
PR 16-SEP-1999; 99US-0154039P
PR 20-SEP-1999; 99US-0154779P
PR 22-SEP-1999; 99US-0155139P
PR 23-SEP-1999; 99US-0155486P
PR 24-SEP-1999; 99US-0155659P
PR 28-SEP-1999; 99US-0156458P
PR 29-SEP-1999; 99US-0156596P
PR 04-OCT-1999; 99US-0157117P
PR 05-OCT-1999; 99US-0157753P
PR 06-OCT-1999; 99US-0157865P
PR 07-OCT-1999; 99US-0158029P
PR 08-OCT-1999; 99US-0158322P
PR 12-OCT-1999; 99US-0158369P
PR 13-OCT-1999; 99US-0159293P
PR 13-OCT-1999; 99US-0159294P
PR 13-OCT-1999; 99US-0159295P
PR 14-OCT-1999; 99US-0159329P
PR 14-OCT-1999; 99US-0159330P
PR 14-OCT-1999; 99US-0159331P
PR 14-OCT-1999; 99US-0159637P
PR 14-OCT-1999; 99US-0159638P
PR 18-OCT-1999; 99US-0159584P
PR 21-OCT-1999; 99US-0160741P
PR 21-OCT-1999; 99US-0160767P
PR 21-OCT-1999; 99US-0160768P
PR 21-OCT-1999; 99US-0160770P
PR 21-OCT-1999; 99US-0160814P
PR 21-OCT-1999; 99US-0160815P
PR 22-OCT-1999; 99US-0160880P
PR 22-OCT-1999; 99US-0160881P
PR 22-OCT-1999; 99US-0160989P
PR 25-OCT-1999; 99US-0161404P
PR 25-OCT-1999; 99US-0161405P
PR 25-OCT-1999; 99US-0161406P
PR 26-OCT-1999; 99US-0161359P
PR 26-OCT-1999; 99US-0161360P
PR 26-OCT-1999; 99US-0161361P
PR 28-OCT-1999; 99US-0161920P
PR 28-OCT-1999; 99US-0161922P
PR 28-OCT-1999; 99US-0161993P
PR 29-OCT-1999; 99US-0162142P

```

Query Match 12.3%; Score 281.5; DB 3; Length 485;
 Best Local Similarity 26.7%; Pred. No. 1.1e-17;
 Matches 109; Conservative 61; Mismatches 166; Indels 73; Gaps 16;

```

QY 37 ILTAGGLACNLSAIGSLIERTEIDPSLEITCYRGVYGLLSDSPYTAVERKKAGVL 96
DB 96 ITCGGLCGMLTAVREVVSSSYMGVRRILIGIDGGYGYAKNTPIINSKYV---VNDI 152
QY 97 QRFSGSVIGNSRHKLTNNVDQVCRGLVKEGEDPOKVAADQLVADGVDLHTIGDDDTNTA 156
DB 153 HRRGGIITISR-----GSHDTNKK-VDSIDRGINGVYIIGSGDGTORG 195
QY 157 AADLAFLARNYGYLTIGLPTVDNDVFPPIKISIGAMTAAEGARYF--NNVVAENNAN 214
DB 196 ASVIFEEIRRRRLKVAVVGIPTKIDNDIPVICKSFGFDPAVEBAQAIYAAHVEAESNEN 255
QY 215 PMLLVHEVMGNCGMLTAATQAEYKLLDRAEMLPBELGLTRESYEVNAVFPBEAIDLE 274

```

```

DB 256 GIGFV--KLMGRYSGYA-----MYATLASRDVDCCLIPESPFYIE 294
QY 275 AEA-----KRLREVMKDCVNIFVSEGAVEAIVAEQAKGSEVPDAFGH1KLDAY 327
DB 295 GEGGLFEFIERLRKDHGMV-----IVALGAGQDLMCKSMES-----TPMDASGNKLLKOV 346
QY 328 NPGKWFSGEQFAMQIAEKLTVQ---KSGYFARASNVDDKRLKSCADLAVECAFRRE 383
DB 347 --GLWLSQSIDHDFKKNKVMNLKXYIDPTYMIKAVPSNASDVMVYCTLLAQSLVHGAM--- 401
QY 364 SGVIGHDD--NGNVLRAIEFPRIKGGKPFNIDT--MFSNLSSEIGP 428
DB 402 AGTYGTGLVNGR-QTYIIPFRIETQNNVITDRMAR-LISSSTNDP 448

```

RESULT 10
 ABU19327
 ID ABU19327 standard; protein; 448 AA.
 AC ABU19327;
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #4854.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Borrelia burgdorferi.
 XX
 PN WO200277183-A2.
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00845242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-029926/02.
 DR N-PsDB; AKA23197.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 47251; 1766bp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 623 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a


```
XX Oryza sativa.
OS Key Location/Qualifiers
FH Misc-difference 130
FT /note= "encoded by codon AAT"
XX
XX MO9505457-A1.
XX
XX 23-FEB-1995.
XX
XX 16-AUG-1994; 94WO-JP001352.
XX
XX 19-AUG-1993; 93JP-00226454.
XX
XX (NISR ) JAPAN TOBACCO INC.
XX
XX Hiyoeshi T, Mine T, Kasaoka K, Tyson HR, Page MJA;
XX
XX WPI, 1995-098757/13.
XX
XX N-PSDB; AAQ85984.
XX
XX DNA coding for fructose-6-phosphate 1-phosphotransferase - of plant
XX origin, for prodn. of transformant plant cells with altered sugar
XX content.
XX
XX Claim 7, Page 46-49; 79pp; Japanese.
XX
XX The amino acid sequence of the novel ATP-dependent fructose-6-phosphate 1
XX -phosphotransferase enzyme (EC 2.7.1.11; PFK) encoded by the Oryza sativa
XX (rice) gene, pPFK-OS, as given in the specification. Plants transformed
XX with genes encoding PFK (see AAQ85982-86) can express the enzyme. The
XX transformed plants can produce varieties that have altered sugar content
XX on storage at low temperatures. (Updated on 25-MAR-2003 to correct PN
XX field.)
XX
XX Sequence 469 AA;
SQ
Query Match 11.8%; Score 269.5; DB 2; Length 469;
Best Local Similarity 26.8%; Pred. No. 1.4e-16;
Matches 109; Conservative 57; Mismatches 172; Indels 69; Gaps 16;
QY 37 ILTAGGLAPCLNSAIGSLIERTEIDPSIEITCYRGYKGLLDGSPVTAVERKKAGVL 96
DB 74 IVTGGGLCPGLNTVIRIETVCGLDNMGVSRVLGIGGVRGYFACTIDLSP---KSVNDI 130
QY 97 QRFSGSVIGNSRVKLTNVKDCYKGLVKEGEDPQKVAADOLVKQGVDIHTIGDPTNTA 156
DB 131 HRRGGTVLGTSR-----GHDYTKI-VDSDRGRINQYVYIGSGDTQRG 173
QY 157 AADLAAPLARNNYGLTVIGLPTVNDVPIKOSLGAWTAAEQARYFMNVVAENNANPR 216
DB 174 AGVIFEEIRRRGLKVAAGVIGIPKTINDIPIDRSFGPTVAEBQRAINAAHVEAGSAEN 233
QY 217 MLIYHEVNGRNGWLTAATAQERYKLDRAEWLPGLGLTRSEYVHAVFPEMAIDLEAE 276
DB 234 GIGLVKLMGRHSGFLA---HYATLASR-----DVDCCILIPSPFLBEG 274
QY 277 A-----KRLAEVNDKDCVNIIFYSEGAGVEAIVAEMQAGQEVPRDAFGH-IKLDAYN 328
DB 275 GGLFYLELGRLEKNGHNV---IYVAGEGGQKLINETKESMG---KDAISGSLILDV-- 324
QY 329 PGMFGEOPAQMIIGAECTLVQ---KSGYFARASASNVDDMLIKSCADLAVECAFRRES 384
DB 325 -GLMISQRIKEHFKKIKTINLKYIDPTVMIRAIPSNAGDNVYCTLLAHVVHGAAGYT 383
QY 385 G-VIHDENDGNVLAIEPRPI--KGGKRPNIDTDMFNSMLSEIQP 428
DB 384 GPTVG--QVNGRHG-YIFPYRIITEKONKVSITDRMMAR-LIASSTNP 426
RESULT 13
AAR71583
```

```
ID AAR71583 standard; protein; 426 AA.
XX
XX AC AAR71583;
XX
XX DT 25-MAR-2003 (revised)
XX DT 12-OCT-1995 (first entry)
XX
XX Raphanus sativus fructose-6-phosphate 1-phosphotransferase.
XX
XX ATP-dependent fructose-6-phosphate 1-phosphotransferase enzyme; plant;
XX potato; Solanum tuberosum; rice; Oryza sativa; maize; Zea mays; radish;
XX Raphanus sativus; Plavaria brownii; primer; expression vector;
XX Agrobacterium tumefaciens; sugar; storage; temperature.
XX
XX Raphanus sativus.
XX
XX MO9505457-A1.
XX
XX 23-FEB-1995.
XX
XX 16-AUG-1994; 94WO-JP001352.
XX
XX 19-AUG-1993; 93JP-00226454.
XX
XX (NISR ) JAPAN TOBACCO INC.
XX
XX Hiyoeshi T, Mine T, Kasaoka K, Tyson HR, Page MJA;
XX
XX WPI, 1995-098757/13.
XX
XX N-PSDB; AAQ85986.
XX
XX DNA coding for fructose-6-phosphate 1-phosphotransferase - of plant
XX origin, for prodn. of transformant plant cells with altered sugar
XX content.
XX
XX Claim 11, Page 58-61; 79pp; Japanese.
XX
XX The amino acid sequence of the novel ATP-dependent fructose-6-phosphate 1
XX -phosphotransferase enzyme (EC 2.7.1.11; PFK) encoded by the Raphanus
XX sativus (radish) gene, pPFK-RS1. Plants transformed with genes encoding
XX PFK (see AAQ85982-86) can express the enzyme. The transformed plants can
XX produce varieties that have altered sugar content on storage at low
XX temperatures. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 426 AA;
SQ
Query Match 11.5%; Score 263; DB 2; Length 426;
Best Local Similarity 25.8%; Pred. No. 5.2e-16;
Matches 108; Conservative 65; Mismatches 173; Indels 72; Gaps 17;
QY 37 ILTAGGLAPCLNSAIGSLIERTEIDPSIEITCYRGYKGLLDGSPVTAVERKKAGVL 96
DB 50 IVTGGGLCPGLNTVIRIETVCGLSVYGVKILIEGGVRGYFARNTIDLD---KTVNDI 106
QY 97 QRFSGSVIGNSRVKLTNVKDCYKGLVKEGEDPQKVAADOLVKQGVDIHTIGDPTNTA 156
DB 107 HRRGGTVLGTSR-----GHDYTKI-VDSDRGRINQYVYIIGSGSGK 149
QY 157 AADLAAPLARNNYGLTVIGLPTVNDVPIKOSLGAWTAAEQARYFMNVVAENNANPR 216
DB 150 AAVIFEEIRRRGLKVAAGVIGIPKTINDIPIDRSFGPTVAEBQRAINAAHVEATSPEN 209
QY 217 MLIYHEVNGRNGWLTAATAQERYKLDRAEWLPGLGLTRSEYVHAVFPEMAIDLEAE 276
DB 210 GIGLVKLMGRSGFLAM---YATLASR-----DVDCCILIPSPFLBEG 250
QY 277 -----AKRLAEVNDKDCVNIIFYSEGAGVEAIVAEMQAGQEVPRDAFGH-IKLDAYN 329
DB 251 GGLFYLELGRLEKNGHNV---IYVAGAG--ODLAEBSNEOSTTL-KDASGKLLQDV-- 302
QY 330 GGMFGEOPAQMIIGAECTLVQ---SGYFARASASNVDDMLIKSCADLAVECAFRRESGV 386
DB 303 GLMISQRIKDHFAKMTINLKYIDPTVMIRAVPSNASD---NVCCITLLAQSAYH---GV 355
```


QY 387 I-GHEDNGNVLA-----IEFRI--KGGKPFNIDTFWNSMLSEIQPKGKVEVSH 437
 DB 356 MAGYNGFTVGLVGRHTYIPFYRIITEKONKVVITDRMMAR-LLSTNQSPFMGDDHH 412

RESULT 14

ID AAR71579 standard; protein; 485 AA.

AC AAR71579;

DT 25-MAR-2003 (revised)
 DT 11-OCT-1995 (first entry)

DE Solanum tuberosum fructose-6-phosphate 1-phosphotransferase.

KM ATP-dependent fructose-6-phosphate 1-phosphotransferase enzyme; plant;
 KM potato; Solanum tuberosum; rice; Oryza sativa; maize; Zea mays; radish;
 KM Raphanus sativus; Flaveria browii; primer; expression vector;
 KM Agrobacterium tumefaciens; sugar; storage; temperature.

OS Solanum tuberosum.

FN WO9505457-A1.

PD 23-FEB-1995.

PF 16-AUG-1994; 94NO-JP001352.

PR 19-AUG-1993; 93JP-00226454.

PA (NISR) JAPAN TOBACCO INC.

PI Hiyeohi T, Mine T, Kasaka K, Tyson HR, Page MD.

DR WPI; 1995-098757/13.

DR N-PSDB; AAQ85982.

PT DNA coding for fructose-6-phosphate 1-phosphotransferase - of plant
 origin, for prodn. of transformant plant cells with altered sugar
 content.

PS Claim 4; Page 33-37; 79pp; Japanese.

CC The amino acid sequence of the novel ATP-dependent fructose-6-phosphate 1
 CC -phosphotransferase enzyme (EC 2.7.1.11; PFK) encoded by the potato
 CC (Solanum tuberosum) gene PFK32. The DNA sequence was isolated from a
 CC potato cDNA library using primers AAQ85987-91. The sequence was used to
 CC construct the plant expression vector PPK(35S) which was transformed
 CC into potato plants (Bintje variety) via Agrobacterium tumefaciens. The
 CC transformed plants can express the enzyme and produce plant varieties
 CC that have altered sugar content on storage at low temperatures. (Updated
 CC on 25-MAR-2003 to correct FN field.)

CC SQ Sequence 485 AA;

Query Match 11.5%; Score 262; DB 2; Length 485;

Best Local Similarity 25.7%; Pred. No. 7.8e-16;

Matches 113; Conservative 56; Mismatches 179; Indels 92; Gaps 17;

QY 7 YHLTADIRCFHMFLLNFNFTLMNPKKVALITAGGLACPLNSAIGSLTERYTEIDPSIE 66

DB 86 YRSSDDVRAC-----IVTCGLCPGLNVIIEIYHSLDYMGNVK 125

QY 67 IICRYGGYGLLIGDSYPTAVSRKAGVQRFSSVIGNSRVKLTNVKDCVRGLVKEG 126

DB 126 VFGIDGGYGFYSKNIINLTP--KTVNDIHKKGCTILGSSR-----CG 166

QY 127 EDPQKVAADQLVVDGDIHTTGDDNTNPAADLAFIARNNGLVIGLPTVDNDVFP 186

DB 167 HDTTKI-VDSIQDREINOVYIIGGDDGQGAAYVEIRRRGLKVIAGIKPTINDIPV 225

QY 187 IKQSLGAMTAAEGAGYFNNVVAENNANPRLMIVHEMGRNCGLTAATAQERYKLLDRA 246

DB 226 IDKSFQFTAVEBAQAIINAAHVEASANGICVWLMGRYSGFIA----- 271

QY 247 EMLPELGLTRESIEYVAIVPEPAIDLEAA-----KRLREVMKQDCNIFVEEGAG 289

DB 272 -----MYATLASADVDLCLIPESPFYLEDGGLFEYIERLKNGMV-----LVIAEGAG 322

QY 300 VEALVAE-MQAKQOEVRDPAFGHIKLDVAVNPCKMFGEOFAQMIGAKETLVQ-----KSG 352

DB 323 QELIAEENMAAKNEQ---DASGNKLIQDV--GLMISQKIRDNH-ARKTKMPTILKXIDPT 376

QY 353 YEPARASAVDDMRILKSCADLAVECAFRRSGVIGHDED--NGNVLRAIEFPRIKGGKP 410

DB 377 YMIRAVPSNASDN---VCTLLAQSCVHGAMAGYTGFTSGLVNGR-QTYIPFRIITEKON 432

QY 411 FNIDTD--WFNSMLSEIQP 428

DB 433 MVVITDRMMAR-LLSTNQP 451

RESULT 15

ID ADC07918 standard; protein; 496 AA.

AC ADC07918;

DT 18-DEC-2003 (first entry)

DE Rice protein sequence Seg ID184 related to grain filling.

CC plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
 CC carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
 CC tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
 CC wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;
 CC gene; de; plant.

OS Oryza sativa.

PN MO2003000905-A2.

PR 22-JUN-2001; 2001US-0300112P.

PR 26-SEP-2001; 2001US-0325277P.

PR 20-DEC-2001; 2001US-0342327P.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;

PI Glazebrook J, Katagiri F, Kreps J, Provart N, Riecke D;

DR N-PSDB; ADC07917.

DR WPI; 2003-229341/22.

PT New plant genes encoding polypeptides having an activity involved in or
 PT associated with the synthesis, metabolism or degradation of carbohydrates
 PT in the plant grain useful in generating plants having improved
 PT nutritional properties.

PS Claim 1; SEQ ID NO 184; 130pp; English.

CC This invention, in the area of plant biotechnology, relates to novel
 CC polynucleotides comprising a nucleotide sequence encoding a protein which
 CC is involved in or associated with the synthesis, metabolism or
 CC degradation of carbohydrates in the plant grain and the expression of
 CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
 CC sugarbeet, wheat, and rice. The invention may be useful for the
 CC improvement of protein, oil, starch, fibre and moisture content of the
 CC cereal grains. In addition, carbohydrate levels may be modified to a more

CC desirable level using the present invention. The present sequence is the
CC amino acid sequence of a rice protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/publishedpct_sequences.

XX
SQ Sequence 496 AA;

Query Match 11.4%; Score 261.5; DB 7; Length 496;
Best Local Similarity 26.2%; Pred. No. 9e-16;
Matches 107; Conservative 54; Mismatches 180; Indels 67; Gaps 13;

```
QY 36 ALITAGGLAPCLNSAIGSLIERYTEIDPSIELICRYGKGLLDGSPYPTAEVRKXGV 95
DB 95 AIVTCGGICPGELNIVRELVCGLDMDYGVTSVVGIEGKGYSRVTALTP--KSVND 151
QY 96 LORFGGSVIGNSRVKLTNVKDCVKRGLVEGEDPQKVAADQLVKDGVDLHTIGGDDTNT 155
DB 152 IHKRGSTVLGTSR-----CGHDTGKI-VDSIKDRGINQVYIIIGDGTQK 194
QY 156 AADLAFLARNNYGLTVIGLPKYDNDVFPFKQSLGAWTAABQGARYPFNVVAENNANP 215
DB 195 GASVIYEVEVRRRGLKCSVGVGPKTIDNDIAVIDKSFQFDTAVEAQRAINAAHVEAESAE 254
QY 216 RMLIVHEVWGRNCGWLTATAQRYKLDRAEMLPBLGLTRESYEHAVFVPEMAIDLEA 275
DB 255 NGIGVVKLMGRNSGFLA-----MYATLASRDVDCCLIPESPPYLEG 295
QY 276 EA-----KRLREVWDKVCVNIIFVSEAGYEAIVAEWQAKGOEYPRDAFGHIKLDVAVN 328
DB 296 KGGLEIEFIEKRLKDNQHNW----IYVABGAGQDLIAKSNPFVDTQ---DASGNKLL--LD 346
QY 329 PKGMFGEOPAQMTGAETL-----VQKSGYFAPASASNTVDMRLIKSCADLAVECAFRE 383
DB 347 VGLWLSQKIKDHPKKRNFPITLKYIDPYMIRAVRASNASN---VYCTLLAHSAHLGAM 403
QY 384 SGVIGHDEDNQVLA--IEFPRI--KGGKPFNIDTDMFMSMLSEIGOP 428
DB 404 AGYTGFTVAPVNGRHAYIPFYRITEKONKRVITTRMMAR-VLCSTNQP 450
```

Search completed: February 24, 2005, 14:19:27
Job time : 176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 14:13:35 / Search time 44 Seconds
(without alignments)
741.401 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTPYHLLTADIRCFHMF.....FNSMLEIGQPKGKGVESH 437

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents, AA:
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6C_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/6D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2284	100.0	437	4	US-09-934-901-16
2	2284	100.0	437	4	US-09-934-868-6
3	2284	100.0	437	4	US-10-321-210-16
4	2284	100.0	437	4	US-10-320-874-16
5	270	11.8	483	2	US-08-416-870C-4
6	264.5	11.6	469	2	US-08-416-870C-6
7	262	11.5	485	2	US-08-416-870C-2
8	261	11.4	436	2	US-08-416-870C-10
9	260	11.4	426	2	US-08-416-870C-10
10	227	9.9	438	4	US-09-489-039A-12258
11	226.5	9.9	345	3	US-09-134-001C-31774
12	220	9.6	330	4	US-09-543-681A-8337
13	220	9.6	335	4	US-09-583-110-5099
14	218.5	9.6	321	4	US-09-134-000C-5332
15	215.5	9.4	323	4	US-09-107-532A-6470
16	211.5	9.3	522	4	US-08-416-870C-8
17	210	9.2	389	4	US-09-949-016-9488
18	206.5	9.0	781	1	US-08-380-690-2
19	204.5	9.0	784	1	US-09-849-016-6251
20	201.5	8.8	798	4	US-09-949-016-11096
21	201	8.8	343	4	US-09-887-054-2
22	182	8.0	307	4	US-09-710-279-2122
23	181	7.9	550	4	US-09-198-452A-225
24	181	7.9	556	4	US-09-438-185A-210
25	179	7.8	556	4	US-09-438-185A-162
26	171.5	7.5	544	4	US-09-198-452A-180
27	148	6.5	696	4	US-09-949-016-8015

28	120	5.3	169	4	US-09-107-433-2681	Sequence 2681, Ap
29	113.5	5.0	845	4	US-09-198-452A-458	Sequence 458, App
30	113.5	5.0	847	4	US-09-438-185A-439	Sequence 439, App
31	107	4.7	810	4	US-09-583-110-4352	Sequence 4352, App
32	107	4.7	816	4	US-09-107-433-4504	Sequence 4504, App
33	106.5	4.7	543	4	US-09-248-796A-17070	Sequence 17070, A
34	101.5	4.4	1427	4	US-09-538-092-1044	Sequence 1044, Ap
35	100.5	4.4	341	4	US-09-248-796A-17051	Sequence 17051, A
36	99.5	4.4	742	4	US-09-489-039A-7998	Sequence 7998, Ap
37	99.5	4.4	1038	3	US-09-541-782-4	Sequence 4, App1
38	99.5	4.4	1038	4	US-09-723-820-4	Sequence 4, App1
39	99.5	4.4	1038	4	US-10-270-085-4	Sequence 4, App1
40	97.5	4.3	413	4	US-09-603-208A-8	Sequence 6, App1
41	97.5	4.3	524	4	US-09-603-208A-6	Sequence 6, App1
42	97.5	4.3	724	4	US-09-489-039A-12100	Sequence 12100, A
43	96.5	4.2	458	4	US-09-489-039A-9184	Sequence 9184, Ap
44	96	4.2	863	4	US-09-328-352-6730	Sequence 6730, Ap
45	95.5	4.2	813	4	US-09-252-991A-30252	Sequence 30252, A

ALIGNMENTS

US-09-934-901-16	Sequence 16, Application US/09934901
Patent No. 655353	
GENERAL INFORMATION:	
APPLICANT: Kotias, Matheos	
APPLICANT: Odum, J. Martin	
APPLICANT: No. 65535310n, Kelley C.	
APPLICANT: Ye, Rick	
TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN	
FILE REFERENCE: C1619 US NA	
CURRENT APPLICATION NUMBER: US/09/934,901	
PRIOR FILING DATE: 2001-08-22	
PRIOR APPLICATION NUMBER: 60/229,906	
PRIOR FILING DATE: September 1, 2000	
NUMBER OF SEQ ID NOS: 20	
SOFTWARE: Microsoft Office 97	
SEQ ID NO 16	
LENGTH: 437	
TYPE: PR	
ORGANISM: METHYLOMONAS SP.	
US-09-934-901-16	
Query Match	100.0%; Score 2284; DB 4; Length 437;
Best Local Similarity	100.0%; Pred. No. 1.6e-215;
Matches 437; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 DVVTPYHLLTADIRCFHMFENFYTLNKKKVAIITAGGLACLNAGISLTERYTE 60
DB	1 DVVTPYHLLTADIRCFHMFENFYTLNKKKVAIITAGGLACLNAGISLTERYTE 60
QY	61 IDPSIEIICRYRGYGLLLGDSYPTAEVRKKAAGVLOFGSGSVIGNSRYKLTNVDCVGR 120
DB	61 IDPSIEIICRYRGYGLLLGDSYPTAEVRKKAAGVLOFGSGSVIGNSRYKLTNVDCVGR 120
QY	121 GLVKEGEDPQKAAADQVLDVGDVLIHTIGDDTNTAAADLAFLARNNGTLVIGPKTV 180
DB	121 GLVKEGEDPQKAAADQVLDVGDVLIHTIGDDTNTAAADLAFLARNNGTLVIGPKTV 180
QY	181 DNDVPRKOSIGAMPAAGAGAYFNNVVAENNANPRMLIVHVGRCNGMLTAATROEYR 240
DB	181 DNDVPRKOSIGAMPAAGAGAYFNNVVAENNANPRMLIVHVGRCNGMLTAATROEYR 240
QY	241 KLLDRAEWLPELGLTRRESYEVAVPEMAIDLEAEAKRLREVMKVCVNFVEGAGV 300
DB	241 KLLDRAEWLPELGLTRRESYEVAVPEMAIDLEAEAKRLREVMKVCVNFVEGAGV 300
QY	301 EAIIVAEWQKGEVPRDAFGHIKLDVANPCKMFGSOPQOMIGAETTLVOKSGYPARASAS 360
DB	301 EAIIVAEWQKGEVPRDAFGHIKLDVANPCKMFGSOPQOMIGAETTLVOKSGYPARASAS 360

```
QY 361 NVDDMRLLKSCADLAVECAFRESGVI GHDENGVLAIEPRRIKGGKPFNIDTDMNS 420
DB 361 NVDDMRLLKSCADLAVECAFRESGVI GHDENGVLAIEPRRIKGGKPFNIDTDMNS 420
QY 421 MLSEIGOPKGGKVEVSH 437
DB 421 MLSEIGOPKGGKVEVSH 437

RESULT 2
US-09-934-868-6
; Sequence 6, Application US/09934868
; Patent No. 6689601
; GENERAL INFORMATION:
; APPLICANT: Kofias, Matheos
; APPLICANT: Odem, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 437
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
; US-09-934-868-6

Query Match 100.0%; Score 2284; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.6e-215;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWTPYHLTADIRCHWFELNPFYTLNKPCKVAIITAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVTWTPYHLTADIRCHWFELNPFYTLNKPCKVAIITAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEIICYRGYKGLLDGSDYPYTAEVRRKAGVLOFGSGSVIGNSRYKLTNVDCVXR 120
DB 61 IDPSIEIICYRGYKGLLDGSDYPYTAEVRRKAGVLOFGSGSVIGNSRYKLTNVDCVXR 120
QY 121 GLVKEGEDPQKVAADQVLDVGDVLIHTIGDDDTNTAAADLAFLARNNYGLTVIGLPTKV 180
DB 121 GLVKEGEDPQKVAADQVLDVGDVLIHTIGDDDTNTAAADLAFLARNNYGLTVIGLPTKV 180
QY 181 DNDVFPRIKOSLGAMTAAEGARFEMNVVAENNANPRMLIVHEVMGRNGMLTAATAOEYR 240
DB 181 DNDVFPRIKOSLGAMTAAEGARFEMNVVAENNANPRMLIVHEVMGRNGMLTAATAOEYR 240
QY 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSEGAGV 300
DB 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSEGAGV 300
QY 301 EAI VAEWQAKQGEVPRDAFGHI KLDVNPGRKMFGEQFAQMI GAETTLVOKSGYFARASAS 360
DB 301 EAI VAEWQAKQGEVPRDAFGHI KLDVNPGRKMFGEQFAQMI GAETTLVOKSGYFARASAS 360
QY 361 NVDDMRLLKSCADLAVECAFRESGVI GHDENGVLAIEPRRIKGGKPFNIDTDMNS 420
DB 361 NVDDMRLLKSCADLAVECAFRESGVI GHDENGVLAIEPRRIKGGKPFNIDTDMNS 420
QY 421 MLSEIGOPKGGKVEVSH 437
DB 421 MLSEIGOPKGGKVEVSH 437

RESULT 3
US-10-321-210-16
; Sequence 16, Application US/10321210
; Patent No. 6767744
```

```
; GENERAL INFORMATION:
; APPLICANT: Kofias, Matheos
; APPLICANT: Odem, J. Martin
; APPLICANT: No. 6767744ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/321,210
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 437
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
; US-10-321-210-16

Query Match 100.0%; Score 2284; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.6e-215;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWTPYHLTADIRCHWFELNPFYTLNKPCKVAIITAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVTWTPYHLTADIRCHWFELNPFYTLNKPCKVAIITAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEIICYRGYKGLLDGSDYPYTAEVRRKAGVLOFGSGSVIGNSRYKLTNVDCVXR 120
DB 61 IDPSIEIICYRGYKGLLDGSDYPYTAEVRRKAGVLOFGSGSVIGNSRYKLTNVDCVXR 120
QY 121 GLVKEGEDPQKVAADQVLDVGDVLIHTIGDDDTNTAAADLAFLARNNYGLTVIGLPTKV 180
DB 121 GLVKEGEDPQKVAADQVLDVGDVLIHTIGDDDTNTAAADLAFLARNNYGLTVIGLPTKV 180
QY 181 DNDVFPRIKOSLGAMTAAEGARFEMNVVAENNANPRMLIVHEVMGRNGMLTAATAOEYR 240
DB 181 DNDVFPRIKOSLGAMTAAEGARFEMNVVAENNANPRMLIVHEVMGRNGMLTAATAOEYR 240
QY 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSEGAGV 300
DB 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSEGAGV 300
QY 301 EAI VAEWQAKQGEVPRDAFGHI KLDVNPGRKMFGEQFAQMI GAETTLVOKSGYFARASAS 360
DB 301 EAI VAEWQAKQGEVPRDAFGHI KLDVNPGRKMFGEQFAQMI GAETTLVOKSGYFARASAS 360
QY 361 NVDDMRLLKSCADLAVECAFRESGVI GHDENGVLAIEPRRIKGGKPFNIDTDMNS 420
DB 361 NVDDMRLLKSCADLAVECAFRESGVI GHDENGVLAIEPRRIKGGKPFNIDTDMNS 420
QY 421 MLSEIGOPKGGKVEVSH 437
DB 421 MLSEIGOPKGGKVEVSH 437

RESULT 4
US-10-320-874-16
; Sequence 16, Application US/10320874
; Patent No. 6773905
; GENERAL INFORMATION:
; APPLICANT: Kofias, Matheos
; APPLICANT: Odem, J. Martin
; APPLICANT: No. 6773905ton, Kelley C.
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/320,874
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
```

PRIOR APPLICATION NUMBER: 60/229,906
 PRIOR FILING DATE: September 1, 2000
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: Microsoft Office 97
 SEQ ID NO: 16
 LENGTH: 437
 TYPE: PRT
 ORGANISM: METHYLOMONAS SP.
 US-10-320-874-16

Query Match 100.0%; Score 2284; DB 4; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1,66-215;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWPHYLTADIRFCMFELNFFYTLNKKPKVAIITAGGLAPCLNSAIGSLIRYTE 60
 DB 1 DVTWPHYLTADIRFCMFELNFFYTLNKKPKVAIITAGGLAPCLNSAIGSLIRYTE 60
 QY 61 IDPSIEIICRGYKGLLDGSDYPTAAYRKAQVLRFGSGVIGNSRVKLTNVKDCVXR 120
 DB 61 IDPSIEIICRGYKGLLDGSDYPTAAYRKAQVLRFGSGVIGNSRVKLTNVKDCVXR 120
 QY 121 GLVKEGEDQKVAADLVKDGVDILHTIGDDTNTAAALAFLANNTGLTYIGLPTV 180
 DB 121 GLVKEGEDQKVAADLVKDGVDILHTIGDDTNTAAALAFLANNTGLTYIGLPTV 180
 QY 181 DNDVFPKISLGAWTAAEQGARYFMNVVANNANPRLIVHEVMGRNCGMLTAATQER 240
 DB 181 DNDVFPKISLGAWTAAEQGARYFMNVVANNANPRLIVHEVMGRNCGMLTAATQER 240
 QY 241 KLLDRAEWLPEIGLTRESYEVHAFVPEMAIDLEBAKRLREVMDKVCNIFVSEAGV 300
 DB 241 KLLDRAEWLPEIGLTRESYEVHAFVPEMAIDLEBAKRLREVMDKVCNIFVSEAGV 300
 QY 301 EAIIVAEWQAKGEVPRDAGHITKLDVANGKMGFGEPAOMIGAEKTLVQSGYFAPASAS 360
 DB 301 EAIIVAEWQAKGEVPRDAGHITKLDVANGKMGFGEPAOMIGAEKTLVQSGYFAPASAS 360
 QY 361 NVDDMKLISCADLAVECAFRESGVIGHDEBNGVLRALIEPRIGKGFENIDTWFNS 420
 DB 361 NVDDMKLISCADLAVECAFRESGVIGHDEBNGVLRALIEPRIGKGFENIDTWFNS 420
 QY 421 MLSEIGQPKGKVEVSH 437
 DB 421 MLSEIGQPKGKVEVSH 437

RESULT 5
 US-08-416-870C-4
 Sequence 4, Application US/08416870C
 Patent No. 5824862

GENERAL INFORMATION:
 APPLICANT: HIYOSHI, TORU
 APPLICANT: MINE, TOSHIKI
 APPLICANT: KASAKURA, KEISUKE
 APPLICANT: TYSON, ROBERT HUM
 APPLICANT: PAGE, ANTHONY MILES JOHN
 TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
 TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTOR
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 STREET: PO BOX 747
 CITY: FALL CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22040-0747
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/416,870C
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MORPHY JR, GERALD M
 REGISTRATION NUMBER: 28,977
 REFERENCE/DOCKET NUMBER: 760-195P (PCT)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 483 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-416-870C-4

Query Match 11.8%; Score 270; DB 2; Length 483;
 Best Local Similarity 27.6%; Pred. No. 16-17;
 Matches 113; Conservative 53; Mismatches 175; Indels 68; Gaps 15;

QY 36 AITNAGGLAPCLNSAIGSLIRYTEIDPSIEIITCYRGYKGLLDGSDYPTAAYRKAQV 95
 DB 95 AITNAGGLAPCLNSAIGSLIRYTEIDPSIEIITCYRGYKGLLDGSDYPTAAYRKAQV 95
 QY 96 LDFGSGVIGNSRVKLTNVKDCVXRGLVKEGEDQKVAADLVKDGVDILHTIGDDTNT 155
 DB 152 IHKGGTIIIGTSR-----GGHDKPKI-VDSIODRGINOVYIIIGSGTOK 194
 QY 156 AAADLAFLARNNGYGLTVGLPKTVNDVFPKISLGAWTAAEQGARYFMNVVANNANP 215
 DB 195 GAIVTQEVRRRGKRAVAVGIPKTIIDNDIPVLDKSGFPTVAEGRALNAHVEKSAE 254
 QY 216 RMLIVHEVMGRNCGMLTAATQERYKLDRAEWLPEIGLTRESYEVHAFVPEMAIDLEA 275
 DB 255 NGIGVVKLMGRYSGFIA-----MYATLASRDVLDCLIPESPFYLEG 295
 QY 276 EA-----KRLREVMDKVCNIFVSEAGVEAIVAEWQAKGEVPRDAGHITKLDVANG 328
 DB 296 EGGLEBYEYERKLDQGHV---IIVAEAGGELLAAEMLK--TSTAQASGKLLHDV- 348
 QY 329 PKGKFEQD---FAOMIGAEKTLVQ-KSGYFAPASASNVDDMKLISCADLAVECAFRE 383
 DB 349 -GLMISDKIKAFKIPMPITLKYIDPYTMRAPSNASD---VYCTLLAQSCYHGM 404
 QY 384 SGVIGHDE--NGVLRALIEPRIGKGFENIDT--WNSMLSEIGQ 428
 DB 405 AGVTGFTSLVNGR-QTYIPFRIRITEKQNNVVITDRMAR-LLSSTNP 451

RESULT 6
 US-08-416-870C-6
 Sequence 6, Application US/08416870C
 Patent No. 5824862

GENERAL INFORMATION:
 APPLICANT: HIYOSHI, TORU
 APPLICANT: MINE, TOSHIKI
 APPLICANT: KASAKURA, KEISUKE
 APPLICANT: TYSON, ROBERT HUM
 APPLICANT: PAGE, ANTHONY MILES JOHN
 TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
 TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTOR
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 STREET: PO BOX 747
 CITY: FALL CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22040-0747

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,870C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-195P (PCT)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-416-870C-6

```

```

Query Match 11.6%; Score 264.5; DB 2; Length 469;
Best Local Similarity 26.8%; Pred. No. 3,4e-17;
Matches 109; Conservative 56; Mismatches 173; Indels 69; Gaps 16;

```

```

QY 37 ILTAGLAFLAPCLNSAIGSLIERYTEIDPSIEIICRGYGLLLGDSYPTAEVRKAGVL 96
DB 74 IYTCGCLCGINTVIREIVCGLNDYGVSRVIGCGYGFACNTIDSP---KSYNDN 130
QY 97 QRFSGSVIGNSVYKLTNVKDCVKGELVKEGEPQKYADQLVKDVLIHTTGDDTNTA 156
DB 131 HRRGGTVIGTSR-----GHDVTMKI-VDSIODRGINGVYVIGGDTORG 173
QY 157 AADLAFLARNNYGLTVIGLPTVNDVFPKISGAMTAAGARFENNVAAENNANR 216
DB 174 AVITEEIRRRRLKVAIVAGIPTIDNDIVIDRSFGFDVAEAGQALNAHVEAGSAEN 233
QY 217 MLIVHVMGRACGWLTAATAOEYRKLDRBWLPELGLTRESYEVHAFVPEMAIDLEAE 276
DB 234 GIGLVKLMGRHSGFLA---HYATLASR-----DVDCCIIPESPFLBEE 274
QY 277 A-----KRLREVNDKYDCVNI FVSEGAVEAIVAEWQAKGEVPRDAFGH- IKLDAYN 328
DB 275 GGLFRLERKLEKNGHMV---IIVAEGAGQKLI NETKESMG---KDSAGNSILLDV-- 324
QY 329 PKMFGGEORFAMIGAEKTLVQ---KSGYFARASASNVDMRLIKSCADLAVECAFRRRS 384
DB 325 -GLWISQKIKERPKKIKITINLKVIDPTVMIRAIPEASDNVYCTLLAHRVHGAAGYT 383
QY 385 G-VIGHDEBNGVNLRAIEFPRI--KGGKPFNIDTDFNSMLSEIQP 428
DB 384 GTTVG--QVNGRHC-VIPEFRI TEKONKYSITDRMMAR-LISSTNQP 426

```

```

RESULT 7
US-08-416-870C-2
; Sequence 2, Application US/08416870C
; Patent No. 5824862
; GENERAL INFORMATION:
; APPLICANT: HIYOSHI, TORU
; APPLICANT: KASAKURA, KEISUKE
; APPLICANT: TYSON, ROBERT HUM
; APPLICANT: PAGE, ANTHONY WILES JOHN
; TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
; TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTO
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP

```

```

; STREET: PO BOX 747
; CITY: FALL CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,870C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-195P (PCT)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-416-870C-2

```

```

Query Match 11.5%; Score 262; DB 2; Length 485;
Best Local Similarity 25.7%; Pred. No. 6,2e-17;
Matches 113; Conservative 56; Mismatches 179; Indels 92; Gaps 17;

```

```

QY 7 YHLTDIRFCWFLNFNYTLNKPVKVAIILTAGLAFLAPCLNSAIGSLIERYTEIDPSIE 66
DB 86 YFSSDVRAC-----IVTCGCLCPGLINTVIREIVHSLDNYGVNKK 125
QY 67 IICRGYKGLLLGDSYPTAEVRKAGVLRFGSGVIGNSRVKLTNVKDCVKGELVKEG 126
DB 126 VFGIDGGRGYRFSKRIINLTP---KTVDIHKRGGTLIGSSR-----GG 166
QY 127 EDPQKVAADQVVKDGVLIHTTGDDTNTAADAFLARNNYGLTVIGLPTVNDVFP 186
DB 167 HDTTKI-VDSIODREINQVYIIGDGTOKGAIVYIEIRRRGLKIVAGIPTIDNDIPV 225
QY 187 IKOSIGAMTAAGARFENNVAAENNANRMLIVHVMGRACGWLTAATAOEYRKLDR 246
DB 226 IDKSGFDVAEAGQALNAHVEAGSAENGIGVVKLMGRYSGFLA----- 271
QY 247 EMLPELGLTRESYEVHAFVPEMAIDLEAE-----KRLREVNDKYDCVNI FVSEGA 299
DB 272 ----MYATLASRDVLDCLIPESPFYLEGDGLFEYIEKRLKENGHMV---IIVAEGAG 322
QY 300 VEALIVAE--MOAKGEVPRDAFGHILDAVNPQKMFGEQFAMIGAEKTLVQ-----KSG 352
DB 323 QELLAEEVNAHAKNEQ--DASGNKLQDV--GLWISQKIRDFH-ATKIMPTLTKYIDPT 376
QY 353 YFARASASNVDMRLIKSCADLAVECAFRRRSVIGHDED--NGVNLRAIEFPRIKGGK 410
DB 377 YMIRAVPNSASDN---VYCTLLAAGSCVGAAGAGYGFSGLVNGR-QTYIPENRITRKON 432
QY 411 FNIDTD--WNSMLSEIQP 428
DB 433 MVVITDRMMAR-LISSTNQP 451

```

```

RESULT 8
US-09-902-540-16136
; Sequence 16136, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.

```

APPLICANT: Slater, Steven C.
 APPLICANT: Wiegand, Roger C.
 TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 FILE REFERENCE: 38-10(15849)B
 CURRENT APPLICATION NUMBER: US/09/902,540
 CURRENT FILING DATE: 2001-07-10
 PRIOR APPLICATION NUMBER: 60/217,883
 PRIOR FILING DATE: 2000-07-10
 NUMBER OF SEQ ID NOS: 16825
 SEQ ID NO 16136
 LENGTH: 345
 TYPE: PRF
 ORGANISM: Myxococcus xanthus
 US-09-902-540-16136

Query Match 11.4%; Score 261; DB 4; Length 345;
 Best Local Similarity 26.0%; Pred. No. 4,6e-17;
 Matches 96; Conservative 66; Mismatches 141; Indels 66; Gaps 16;

QY 34 KVALITAGGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGSPYTAEVKKA 93
 DB 2 KVALITAGGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGSPYTAEVKKA 93
 QY 94 GYLORGGSGVIGNSRV---KLTIVKDCYKGLYKEGDEPOKAAADLVKDVLDIHTIG 150
 DB 58 GLHR-GGTLTGSRVNPFEVGLERVKRA-----IERGHAIVAIIG 101
 QY 151 DDTNTAAALAAFLANNYGLTVIGLPTVNDVPPIKOSLGAWTAEGGARF--MNV 208
 DB 102 EGTLSAATM-----SQEGLRIVGPKIIDNDINATDFPGPDVIAVATEIDRLHST 155
 QY 209 AENNANPMLIVHEVNGRCWMLTAATAOEYKRLDRAEWLPGLTRSEYVHAVFVE 268
 DB 156 AESH---KVIYCEVNGRHVGMAT-----YAGTAGAD-----VLVE 192
 QY 269 MAIDLEAEAKRL--REYMDKVCNIFVSEGAGVEAIVAEOMAKGQEVPRDAFGHKLDA 326
 DB 193 IPADLAKVAEHIOHRAAGRTFSI-VVAEGTRIKLSADQOEQLVTSGLDEAGRRLIG 251
 QY 327 VNGKMGFQFOAGMIGAEKTLVQSGYPARASASNVDMRL-----IKGADLAVECAPR 381
 DB 252 V--GTLAEIERRTFE--TRVSVLGHIOHGAFTAHDRVLAIRYVHAC--DMVANGERG 307
 QY 382 RESGVIGHD 390
 DB 308 KMAALRGND 316

RESULT 9
 US-08-416-870C-10
 Sequence 10, Application US/08416870C
 Patent No. 5824862
 GENERAL INFORMATION:
 APPLICANT: HIYOSHI, TORU
 APPLICANT: MINE, TOSHIKI
 APPLICANT: KASAOKA, KEISUKE
 APPLICANT: TYSON, ROBERT HUM
 APPLICANT: PAGE, ANTHONY MILES JOHN
 TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
 TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTO
 TITLE OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDE
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 STREET: PO BOX 747
 CITY: PAUL CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22040-0747
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/416,870C
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURPHY JR, GERALD M
 REGISTRATION NUMBER: 28,977
 REFERENCE/DOCKET NUMBER: 760-195P (PCT)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 426 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-416-870C-10

Query Match 11.4%; Score 260; DB 2; Length 426;
 Best Local Similarity 25.6%; Pred. No. 8e-17;
 Matches 107; Conservative 65; Mismatches 174; Indels 72; Gaps 16;

QY 37 ILTAGGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGSPYTAEVKKA 96
 DB 50 ITCGGLCPGLNTVIREIVGSLSYMGVYKKGIEGSGFYARNITDLDL--KTVNDI 106
 QY 97 QRFSGSVINSRVKLTIVKDCYKGLYKEGDEPOKAAADLVKDVLDIHTIGDDTNTA 156
 DB 107 HRRGGTILIGTSR-----GSHDTTKI-VTSIDRGINOVYIIGGDSOSKG 149
 QY 157 AADLAFLARNYGLTVIGLPTVNDVPPIKOSLGAWTAEGGARFPMNVVAENNANR 216
 DB 150 AAVTPEIRRGKAVAGIPKTIIDNDIPIDRSRGTDAVEBAQATAAHHVATSFEN 209
 QY 217 MEIVHEVNGRCWMLTAATAOEYKRLDRAEWLPGLTRSEYVHAVFVEPAIDLEAE 276
 DB 210 GIGLVKLMRGYSGFIA-----MYATLASRDVDCCLIPSPFFLEGG 250
 QY 277 -----AKRLAEVNDKVCNIFVSEGAGVEAIVAEOMAKGQEVPRDAFGHKLDAVNP 329
 DB 251 GGLFEFIRGLKEIGHMV---IYIAEGAG--ODLAESENGSTTL--KDSGNKLLQDV-- 302
 QY 330 GKMFGEQFOAGMIGAEKTLVQ---SGYFARASASNVDMRLIKSCADLAVECAFRESGV 386
 DB 303 GLMISORIKDHPAKKOTLNLKTIIDPTMIRAVPSNAD---NVCTTLAQSAYH---GV 355
 QY 387 I-GHDEDNQNVLA---IEFPRI--KGGKPNIDTDFNSMLSEIGQKGGKVEVSH 437
 DB 356 MGVNGFTVGLVNGRHTYIPFYRIETEKQKVVITDBMMAR--LLSSTNGSPFMKHDH 412

RESULT 10
 US-09-489-039A-12258
 Sequence 12258, Application US/09489039A
 Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gary Breton et. al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709,2004001
 CURRENT APPLICATION NUMBER: US/09/489,039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747
 PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 12258
 LENGTH: 338
 TYPE: PRF
 ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-12258

```

Query Match      9.9%; Score 227; DB 4; Length 338;
Best Local Similarity 23.8%; Pred. No. 9.8e-14;
Matches 102; Conservative 63; Mismatches 138; Indels 126; Gaps 19;

OY 13 IRCHWFPLNPFYTLNKKKVAITTAGGLAPCLNSAIGSLIERYTEIDPSIEIICYRG 72
DB 8 IIFC---FLSSEVVM-----KKIGVITSGDAPGNNAAIRGVYR--AALTEGLEVPGIYD 58
OY 73 GYKGL-----LLGDSYPTAEVRKKAQVLQRFSGSVIGSRVYKLTNVKDCVKGGLVKEGE 127
DB 59 GYLGLEYEDRMVOLDRYSVSDMINR-----GGTFLGSR--PPEFRE-----E 98
OY 128 DQKVAADQVLDVGDVLIHTIGDDTNTAADAFLAANNGLVYIGPKTVNDVFPFI 187
DB 99 HIRAAVIEKKRKGDLALVIGDGSYGMARL-----TEMGFCPTGLPGTIINDIKCT 152
OY 188 KOSLGAMTAEOGARFNNVVAENNANPRLIYHEVGNRCGMLTAATQOERYKLLDRAE 247
DB 153 DYTIGFTLSTVVAIDRLRDTSSSHQHSIV--EVMGYCCDDLTAAM-----IAGGCE 206
OY 248 WL--DELGLTRESYEVAHVPEMAIDLEBAK-----RLREVMKVDQVNI 293
DB 207 FIMVEFEVETRD-----DLVAEIKAGIAKKGKHAIVATEHMCDDVDLAS 252
OY 294 VEEGAGVE--AIVAMQAKGOEVRPDAFGHITKDAVNGKKRFGEPFQMIGAEKTVOK 350
DB 253 IEKETGERETRAVTLGHIGQSGSPVPYDRI-----LASRMGAVALIELL 295
OY 351 SGYFARASANDVDMRLIKSCADLAVECAFRESGVIGHDEDNVLAIEFPRIKGRKP 410
DB 296 QOHGR-----CYGIGNEKLVHND-----IIDAIEMK-----RP 325
OY 411 FNIDTFWEN 419
DB 326 FK--NDWLD 332

RESULT 11
US-09-134-001C-3774
; Sequence 3774, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3774
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3774

Query Match      9.9%; Score 226.5; DB 3; Length 345;
Best Local Similarity 25.4%; Pred. No. 1.1e-13;
Matches 105; Conservative 68; Mismatches 136; Indels 105; Gaps 19;

OY 26 YTLNKKPKKVAITTAGGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLGDSPYV 85
DB 21 YVVM---KXIAVLTSGGDSPGNNAVRAVTR--TAIVNNIEVYGVYQYGLDDDIHKL 75
OY 86 TAEVRKKAQVLQRFSGSVIGSRVYKLTNVKDCVKGGLVKEGDPQKVAADQVKGQVDIL 145
DB 76 --ELDSVGDIIOR--GGTFLEFSAR-----CPQ--FKE--EDVRKKAIEENRKKGIEGL 120
OY 146 HTIGGDDTNTAADAFLAARNNGVLTGLPKTVNDVFPFIKOSIGAMTAEOGARFYM 205

```

```

DB 121 VVIGDGSYRGAQRISSECKE-----IOTIGIPGTINDINGDFTIGFDTA-----L 168
OY 206 NVVAEN-----NANPRLIYHEVGNRCGMLTAATAOERYKLLDRAEWLPELGLTRE 257
DB 169 NTIISVDKIRDTASSHARTFLV--EVMGRDCG-----DLALW---AGL--- 207
OY 258 SYEVNAVVPMAIDLEBAKRLREVMKVDQVNI FVEGAGVEAIVAMQAKGOEVRPD 317
DB 208 SVGAETIVLPEVNTDICKOVAEKIEGIRGKKHSI-----VVAEGCMSGGECADE 258
OY 318 AFGHITLDAVNGKWFGBQFQMIGAEKTVOKSGYFARASANDVDMRLIKSCADLAVE 377
DB 259 LKTYINID-----TRYSVLGHIGQSGSPSCADVLASRLGAYVE 298
OY 378 CAFRES-----GVIGHDEDNVLAIEFPRIKGRKPNIDTFWENSLSEIQ 427
DB 299 LKQETAKAGVI-----RNNQLTSTPFDI-----PAESDRKKNQWYELAK 341

RESULT 12
US-09-543-681A-8337
; Sequence 8337, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 8337
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-8337

Query Match      9.6%; Score 220; DB 4; Length 330;
Best Local Similarity 28.9%; Pred. No. 4.6e-13;
Matches 81; Conservative 40; Mismatches 95; Indels 64; Gaps 10;

OY 28 LMKPKKVAITTAGGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLGDSPYPTA 87
DB 6 MWNGIKRIGVLTSGDAPGNNAAIRGVYR--AALSEGLEVYGIFDGYMGLY----- 54
OY 86 EVRKAQVLQRF-----GGSVIGNR---VKLTNVKDCVKGGLVKEGDPQKVAADQ 136
DB 55 ENRMKK--LDRFSVSDMINRGOTFLGSARFPPEFREDNR-----AVALEN 97
OY 137 LVKQGVDIHTIGDDTNTAADAFLAARNNGVLTGLPKTVNDVFPRIKOSIGAMTA 196
DB 98 MKQNELDALVIGDGSYLGAKKL-----TEAGFPCGLPETINDVAGTYTIGYFTA 151
OY 197 ABQGARFNNVVAENNANPRLIYHEVGNRCGMLTAATA-----QERYKLLDRA 246
DB 152 LETAVEAIDRLRDTSTSHKRISIV--EVMGYCCDDLTAALAGGCEPVVLPESSELPFRD 210
OY 247 EMLPELGLTRESYEVAHV-----VPEMAIDLEBAK 278
DB 211 ELAETIKAGIRGKHAIVATEHVCDVHEIARFTEAETK 250

RESULT 13
US-09-583-110-5099
; Sequence 5099, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATHO0-07A

```


/ CURRENT APPLICATION NUMBER: US/09/583,110
 / CURRENT FILING DATE: 2000-05-26
 / PRIOR APPLICATION NUMBER: US/09/107,433
 / PRIOR FILING DATE: 1998-06-30
 / PRIOR APPLICATION NUMBER: US 60/085,131
 / PRIOR FILING DATE: 1998-05-12
 / PRIOR APPLICATION NUMBER: US 60/051,553
 / PRIOR FILING DATE: 1997-07-02
 / NUMBER OF SEQ ID NOS: 5322
 / SEQ ID NO 5099
 / LENGTH: 335
 / TYPE: PR1
 / ORGANISM: Streptococcus pneumoniae
 / US-09-583-110-5099

Query Match 9.6%; Score 220; DB 4; Length 335;

Best Local Similarity 25.9%; Pred. No. 4,7e-13; Matches 88; Conservative 55; Mismatches 115; Indels 82; Gaps 16;

QY 33 KVAAILTAGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGSPYTAVERKK 92
 DB 2 KRIAVLTSGDAPGMAAIRAVRQ--AISEGMEVFGIYDYGAVGAGIHPDA---AS 56
 QY 93 AGVLFQFGSVIGNSR---VKLTNVKDCVKGGLVEGEDPOKVAADOLVKQVDILHTI 148
 DB 57 VGDIIIRGGTFLHSARYPEFAQL-----EG---QLKGIEQLKKIGIEGVVI 100
 QY 149 GDDDTTAADLAFLARNNYGLTVIGLPTVNDVFPKISGATATAEGARYFMNV 208
 DB 101 GGDGSHGAMRL-----TEHGPAIGLGTIDNDVGTDFPTGPTAVTT--AMDALNKI 153
 QY 209 AENNAFPMILVHEVWNGC---WLTATAQEVKRLDRAEWLDELGTRE----- 257
 DB 154 RRTSSHRRTFVLEVGRNAGDIALWAGIATGAD--EII-----IEAGFKMEDIVASTIK 206
 QY 258 -SYEV--HAVFYPEAIDLEAE-AKRLREVMKDVCNIFVSEGAVEAIVAEWQAKQ 312
 DB 207 AGYECCKKNIIVLAEGVMSAAEFQCKLKEAGDTS---LRVTE-----LGHIGRGS 256
 QY 313 EVPRDAFGHIKLDVNPFGKFGQFQMGAEKTLVQKSG 352
 DB 257 PTARD-----RVLASRMGAHAHVKLKEG 279

RESULT 14

/ US-09-134-000C-5332
 / Sequence 5332; Application US/09134000C
 / Patent No. 6617156
 / GENERAL INFORMATION:
 / APPLICANT: Lynn Doucette-Stamm et al
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 / FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 / FILE REFERENCE: 032796-032
 / CURRENT APPLICATION NUMBER: US/09/134,000C
 / CURRENT FILING DATE: 1998-08-13
 / PRIOR APPLICATION NUMBER: US 60/055,778
 / PRIOR FILING DATE: 1997-08-15
 / NUMBER OF SEQ ID NOS: 6812
 / SOFTWARE: Patent version 3.1
 / SEQ ID NO 5332
 / LENGTH: 321
 / TYPE: PR1
 / ORGANISM: Enterococcus faecalis
 / US-09-134-000C-5332

Query Match 9.6%; Score 218.5; DB 4; Length 321;

Best Local Similarity 24.5%; Pred. No. 6.2e-13; Matches 92; Conservative 54; Mismatches 113; Indels 117; Gaps 15;

QY 33 KVAAILTAGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGSPYTAVERKK 92
 DB 3 KRIAVLTSGDAPGMAAIRAVRQ--AISEGMEVFGIYDYGAVGAGIHPDA---AS 56

QY 93 AGVLFQFGSVIGNSRVLKLTNVKDCVKG-----LVKEGEDPOKVAADOLVKD 140
 DB 52 -----RLDVAVDGDKIIRGGTFLYSARYPEFAITEG---QLKGIEQLKKF 93
 QY 141 GVDILHTITGGDDTTRAADLAFLARNNYGLTVIGLPTVNDVFPKISGLGAWTAABEG 200
 DB 94 GIEGLVVIIGGDSYHGA-----MLYTKRFPVAVGIPGTIDNDIPTDPTGPTA--- 143
 QY 201 ARYFMNVVAEN-----NANPMLVHEVWNGCMTAATAQEVKRLDRAEWLDEL 252
 DB 144 ----INTVLESIDRLDRTATSHVTRFVI-EVNGRNG-----DIALMSGVA 184
 QY 253 GLTRSEYEVHAFVPEMAIDLEAEAKRLREVM--KVCNIFVSEGAVEAIVAEWQAK 310
 DB 185 GGADE-----IIEPHDFDMGNVAKRIEGRDRGKCHL-IILAE----- 224
 QY 311 GGEVPRDAFGHIKLDVNPFGKFGQFQMGAEKTLVQKSGFAPASASNDVRLIKS 370
 DB 225 -----VWGNFPAKLSE-YDFHTRVSIILGHVNGGSPSARDVLAASK 267
 QY 371 CADLAVECAFRESGV 386
 DB 268 FGSYAVELLKEGKGL 283

RESULT 15

/ US-09-107-532A-6470
 / Sequence 6470; Application US/09107532A
 / Patent No. 6583275
 / GENERAL INFORMATION:
 / APPLICANT: Lynn A Doucette-Stamm and David Bush
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 / ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 / NUMBER OF SEQUENCES: 7310
 / CORRESPONDENCE ADDRESS:
 / ADDRESSES: GENOME THERAPEUTICS CORPORATION
 / STREET: 100 Beaver Street
 / CITY: Waltham
 / STATE: Massachusetts
 / COUNTRY: USA
 / ZIP: 02354
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: CD-ROM ISO9660
 / COMPUTER: PC
 / OPERATING SYSTEM: <Unknown>
 / SOFTWARE: ASCII
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/107,532A
 / FILING DATE: 30-Jun-1998
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: 60/085,598
 / FILING DATE: 14 May 1998
 / APPLICATION NUMBER: 60/051571
 / FILING DATE: July 2, 1997
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Ariniello, Pamela Deneke
 / REGISTRATION NUMBER: 40,489
 / REFERENCE/DOCKET NUMBER: GTC-012
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (781)893-5007
 / TELEFAX: (781)893-8277
 / INFORMATION FOR SEQ ID NO: 6470:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 323 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / HYPOTHETICAL: YES
 / ORIGINAL SOURCE:
 / ORGANISM: Enterococcus faecium
 / FEATURES:
 / NAME/KEY: m1sc_feature
 / LOCATION: (B) LOCATION 1...323

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 14:22:36 ; Search time 133 Seconds
(without alignment)
1075.221 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTMPHYLTADIRCHWF.....FNSMLSEIGPKGKGVESH 437

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2284	100.0	437	9	US-09-934-901-16
2	2284	100.0	437	9	US-09-934-868-6
3	2284	100.0	437	10	US-09-941-947A-2
4	2284	100.0	437	14	US-10-320-924-16
5	2284	100.0	437	14	US-10-320-874-16
6	2284	100.0	437	15	US-10-363-567-2
7	2284	100.0	437	16	US-10-321-210-16
8	659	28.9	184	15	US-10-369-493-1348
9	306.5	13.4	350	14	US-10-369-493-8108
10	294	12.9	342	14	US-10-156-761-13610
11	293	12.8	414	15	US-10-425-114-51700
12	293	12.8	525	15	US-10-425-114-44512
13	293	12.8	527	15	US-10-424-599-283989

14	293	12.8	527	15	US-10-425-114-49830	Sequence 49830, A
15	284.5	12.5	341	14	US-10-156-761-14652	Sequence 14652, A
16	279.5	12.2	603	16	US-10-437-963-204464	Sequence 204464, A
17	270	11.8	448	15	US-10-282-122A-47251	Sequence 47251, A
18	269.5	11.8	341	14	US-10-156-761-10359	Sequence 10359, A
19	269.5	11.8	530	16	US-10-767-701-46418	Sequence 46418, A
20	263.5	11.5	769	16	US-10-437-963-114757	Sequence 114757, A
21	261	11.4	345	15	US-10-369-493-19261	Sequence 19261, A
22	259	11.3	541	16	US-10-437-963-108835	Sequence 108835, A
23	255	11.2	434	15	US-10-425-114-43593	Sequence 43593, A
24	254	11.1	507	15	US-10-424-599-271108	Sequence 271108, A
25	253	11.1	319	15	US-10-369-493-2905	Sequence 2905, A
26	250.5	11.0	357	15	US-10-369-493-3669	Sequence 3669, A
27	248	10.9	447	15	US-10-424-599-211067	Sequence 211067, A
28	246.5	10.8	987	15	US-10-369-493-21998	Sequence 21998, A
29	246.5	10.8	987	16	US-10-477-369-41	Sequence 41, Appl
30	241.5	10.6	428	15	US-10-425-114-52834	Sequence 52834, A
31	241.5	10.6	481	15	US-10-424-599-258809	Sequence 258809, A
32	240.5	10.5	320	9	US-09-815-242-15785	Sequence 15785, A
33	240.5	10.5	320	15	US-10-282-122A-72988	Sequence 72988, A
34	240.5	10.5	320	15	US-10-282-122A-75033	Sequence 75033, A
35	240.5	10.5	320	15	US-10-282-122A-76025	Sequence 76025, A
36	239.5	10.5	318	15	US-10-369-493-9652	Sequence 9652, A
37	238.5	10.4	320	9	US-09-815-242-10422	Sequence 10422, A
38	238.5	10.4	320	15	US-10-369-493-23638	Sequence 23638, A
39	238.5	10.4	320	15	US-10-282-122A-56771	Sequence 56771, A
40	235	10.3	352	15	US-10-425-114-44609	Sequence 44609, A
41	234.5	10.3	327	15	US-10-282-122A-78090	Sequence 78090, A
42	232	10.2	461	15	US-10-282-122A-76398	Sequence 76398, A
43	229.5	10.0	295	15	US-10-369-493-8855	Sequence 8855, A
44	229.5	10.0	320	15	US-10-282-122A-56380	Sequence 56380, A
45	227.5	10.0	449	15	US-10-369-493-8876	Sequence 8876, A

ALIGNMENTS

US-09-934-901-16	Sequence 16, Appl	US-09-934-901-16	Sequence 16, Appl
Patent No. US20020110885A1	GENERAL INFORMATION:	Patent No. US20020110885A1	GENERAL INFORMATION:
APPLICANT: Kofas, Matheos	APPLICANT: Kofas, Matheos	APPLICANT: Kofas, Matheos	APPLICANT: Kofas, Matheos
APPLICANT: Odom, J. Martin	APPLICANT: Odom, J. Martin	APPLICANT: Odom, J. Martin	APPLICANT: Odom, J. Martin
APPLICANT: No. US20020110885A1, Kelley C.	APPLICANT: No. US20020110885A1, Kelley C.	APPLICANT: No. US20020110885A1, Kelley C.	APPLICANT: No. US20020110885A1, Kelley C.
TITLE OF INVENTION: IDENTIFYING METHANOTROPHIC BACTERIAL STRAIN	TITLE OF INVENTION: IDENTIFYING METHANOTROPHIC BACTERIAL STRAIN	TITLE OF INVENTION: IDENTIFYING METHANOTROPHIC BACTERIAL STRAIN	TITLE OF INVENTION: IDENTIFYING METHANOTROPHIC BACTERIAL STRAIN
FILE REFERENCE: CL619 US NA	FILE REFERENCE: CL619 US NA	FILE REFERENCE: CL619 US NA	FILE REFERENCE: CL619 US NA
CURRENT APPLICATION NUMBER: US/09/934,901	CURRENT APPLICATION NUMBER: US/09/934,901	CURRENT APPLICATION NUMBER: US/09/934,901	CURRENT APPLICATION NUMBER: US/09/934,901
PRIOR FILING DATE: 2001-08-22	PRIOR FILING DATE: 2001-08-22	PRIOR FILING DATE: 2001-08-22	PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229,906	PRIOR APPLICATION NUMBER: 60/229,906	PRIOR APPLICATION NUMBER: 60/229,906	PRIOR APPLICATION NUMBER: 60/229,906
NUMBER OF SEQ ID NOS: 20	NUMBER OF SEQ ID NOS: 20	NUMBER OF SEQ ID NOS: 20	NUMBER OF SEQ ID NOS: 20
SOFTWARE: Microsoft Office 97	SOFTWARE: Microsoft Office 97	SOFTWARE: Microsoft Office 97	SOFTWARE: Microsoft Office 97
SEQ ID NO 16	SEQ ID NO 16	SEQ ID NO 16	SEQ ID NO 16
LENGTH: 437	LENGTH: 437	LENGTH: 437	LENGTH: 437
TYPE: PRT	TYPE: PRT	TYPE: PRT	TYPE: PRT
ORGANISM: METHYLOMONAS SP.	ORGANISM: METHYLOMONAS SP.	ORGANISM: METHYLOMONAS SP.	ORGANISM: METHYLOMONAS SP.
US-09-934-901-16	Query Match	US-09-934-901-16	Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-207;	Best Local Similarity 100.0%; Pred. No. 4.8e-207;	Best Local Similarity 100.0%; Pred. No. 4.8e-207;	Best Local Similarity 100.0%; Pred. No. 4.8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 DVVTMPHYLTADIRCHWF.....FNSMLSEIGPKGKGVESH 60	QY	1 DVVTMPHYLTADIRCHWF.....FNSMLSEIGPKGKGVESH 60
DB	1 DVVTMPHYLTADIRCHWF.....FNSMLSEIGPKGKGVESH 60	DB	1 DVVTMPHYLTADIRCHWF.....FNSMLSEIGPKGKGVESH 60
QY	61 IPPSIEIICRGYKGLLDSDVPYTAEVKKAAGVLOFGSGSVIGNSRKLTVNKCVR 120	QY	61 IPPSIEIICRGYKGLLDSDVPYTAEVKKAAGVLOFGSGSVIGNSRKLTVNKCVR 120
DB	61 IPPSIEIICRGYKGLLDSDVPYTAEVKKAAGVLOFGSGSVIGNSRKLTVNKCVR 120	DB	61 IPPSIEIICRGYKGLLDSDVPYTAEVKKAAGVLOFGSGSVIGNSRKLTVNKCVR 120
QY	121 GLVKGEDPQKVAADQVLDVGVDIHTTGGDDTNTAAADLAFLANNYGLTVIGLPKTV 180	QY	121 GLVKGEDPQKVAADQVLDVGVDIHTTGGDDTNTAAADLAFLANNYGLTVIGLPKTV 180

DB 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV 180
QY 181 DNDVFPKISLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGLTAATAOEYR 240
DB 181 DNDVFPKISLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGLTAATAOEYR 240
QY 241 KLLDRAEWLPELGLTRRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSEGAGV 300
DB 241 KLLDRAEWLPELGLTRRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSEGAGV 300
QY 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKWFGEQFAOMIGAETLVOKSGYFARASAS 360
DB 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKWFGEQFAOMIGAETLVOKSGYFARASAS 360
QY 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEDNQVLAIEPRRIKGGKPFNIDTDFNS 420
DB 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEDNQVLAIEPRRIKGGKPFNIDTDFNS 420
QY 421 MLSEIGQPKGKVEVSH 437
DB 421 MLSEIGQPKGKVEVSH 437

RESULT 2

US-09-934-868-6
; Sequence 6, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 437
; TYPE: PRF
; ORGANISM: METHYLIOMONAS SP.
US-09-934-868-6

Query Match 100.0%; Score 2284; DB 9; Length 437;
Best Local Similarity 100.0%; Pred. No. 4.8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWTPYHLTADIRGCHWFPLNFNFTYTLNNKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVTWTPYHLTADIRGCHWFPLNFNFTYTLNNKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEIICRGYGYGLLGDSPYPTAEVRKKAQVLOFRGSGVIGNSRVKLTNVDCVVR 120
DB 61 IDPSIEIICRGYGYGLLGDSPYPTAEVRKKAQVLOFRGSGVIGNSRVKLTNVDCVVR 120
QY 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV 180
DB 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV 180
QY 181 DNDVFPKISLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGLTAATAOEYR 240
DB 181 DNDVFPKISLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGLTAATAOEYR 240
QY 241 KLLDRAEWLPELGLTRRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSEGAGV 300
DB 241 KLLDRAEWLPELGLTRRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSEGAGV 300
QY 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKWFGEQFAOMIGAETLVOKSGYFARASAS 360
DB 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKWFGEQFAOMIGAETLVOKSGYFARASAS 360

QY 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEDNQVLAIEPRRIKGGKPFNIDTDFNS 420
DB 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEDNQVLAIEPRRIKGGKPFNIDTDFNS 420
QY 421 MLSEIGQPKGKVEVSH 437
DB 421 MLSEIGQPKGKVEVSH 437

RESULT 3

US-09-941-947A-2
; Sequence 2, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
; APPLICANT: Brzostowicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: DiCosimo, Deana J.
; APPLICANT: Koffas, Mattheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odum, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Roviatiere, Pierre E.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: CL1903 US NA
; CURRENT APPLICATION NUMBER: US/09/941,947A
; PRIOR FILING DATE: 2001-09-01
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 437
; TYPE: PRF
; ORGANISM: Methyliomonas 16a
US-09-941-947A-2

Query Match 100.0%; Score 2284; DB 10; Length 437;
Best Local Similarity 100.0%; Pred. No. 4.8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWTPYHLTADIRGCHWFPLNFNFTYTLNNKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVTWTPYHLTADIRGCHWFPLNFNFTYTLNNKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEIICRGYGYGLLGDSPYPTAEVRKKAQVLOFRGSGVIGNSRVKLTNVDCVVR 120
DB 61 IDPSIEIICRGYGYGLLGDSPYPTAEVRKKAQVLOFRGSGVIGNSRVKLTNVDCVVR 120
QY 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV 180
DB 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV 180
QY 181 DNDVFPKISLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGLTAATAOEYR 240
DB 181 DNDVFPKISLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGLTAATAOEYR 240
QY 241 KLLDRAEWLPELGLTRRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSEGAGV 300
DB 241 KLLDRAEWLPELGLTRRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSEGAGV 300
QY 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKWFGEQFAOMIGAETLVOKSGYFARASAS 360
DB 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKWFGEQFAOMIGAETLVOKSGYFARASAS 360
QY 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEDNQVLAIEPRRIKGGKPFNIDTDFNS 420
DB 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEDNQVLAIEPRRIKGGKPFNIDTDFNS 420
QY 421 MLSEIGQPKGKVEVSH 437
DB 421 MLSEIGQPKGKVEVSH 437

Db 421 MUSEIQPKGKVEVSH 437

RESULT 4

US-10-320-924-16

Sequence 16, Application US/10320924

Publication No. US20030129721A1

GENERAL INFORMATION:

APPLICANT: Koffas, Matheos

APPLICANT: Odom, J. Martin

APPLICANT: No. US20030129721A1ton, Kelley C.

APPLICANT: Ye, Rick

TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN

FILE REFERENCE: CL1619 US NA

CURRENT APPLICATION NUMBER: US/10/320,924

CURRENT FILING DATE: 2002-12-17

PRIOR APPLICATION NUMBER: US/09/934,901

PRIOR FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/229,906

PRIOR FILING DATE: September 1, 2000

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Microsoft Office 97

SEQ ID NO 16

LENGTH: 437

TYPE: PRT

ORGANISM: METHYLOMONAS SP.

US-10-320-924-16

Query Match 100.0%; Score 2284; DB 14; Length 437;

Best Local Similarity 100.0%; Pred. No. 4.8e-207;

Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVTWPHLTADIRFCHEFFLNFTLNKPKKVAIITAGGLAPCLNSAIGSLIRYTE 60

Db 1 DVTWPHLTADIRFCHEFFLNFTLNKPKKVAIITAGGLAPCLNSAIGSLIRYTE 60

Qy 61 IDPSIEIICRYGGYKGLLDGSPYTAEVKKAGVLQRFSGSVIGNSRYKLTNVKDCVXR 120

Db 61 IDPSIEIICRYGGYKGLLDGSPYTAEVKKAGVLQRFSGSVIGNSRYKLTNVKDCVXR 120

Qy 121 GLVKEGEDPQKVAADLVKDGVDIILHTIGDDTNTAAADLAFLANNNGLTIVIGPKTV 180

Db 121 GLVKEGEDPQKVAADLVKDGVDIILHTIGDDTNTAAADLAFLANNNGLTIVIGPKTV 180

Qy 181 DNDVPEIKOSLGAMTAEQAGARYFMNVVAENNANPMLIVHEVMGNCMLTAATQOER 240

Db 181 DNDVPEIKOSLGAMTAEQAGARYFMNVVAENNANPMLIVHEVMGNCMLTAATQOER 240

Qy 241 KLLDRAEWLPELGLTRESYEVAHVFPVEMAIIDLEAKRLREVMDKVDVCNIFVSEGAGV 300

Db 241 KLLDRAEWLPELGLTRESYEVAHVFPVEMAIIDLEAKRLREVMDKVDVCNIFVSEGAGV 300

Qy 301 EAIIVAMQAKGQEVPRDAFGHIKLDVNPCKMFGQFAOMIGAEKTLVOKSGYFAPASAS 360

Db 301 EAIIVAMQAKGQEVPRDAFGHIKLDVNPCKMFGQFAOMIGAEKTLVOKSGYFAPASAS 360

Qy 361 NVDDMLIKSCADLAIECAFRRRESGYIGHEDNGNVLRAIEFPRIKGGKPFNIDTWMFS 420

Db 361 NVDDMLIKSCADLAIECAFRRRESGYIGHEDNGNVLRAIEFPRIKGGKPFNIDTWMFS 420

Qy 421 MUSEIQPKGKVEVSH 437

Db 421 MUSEIQPKGKVEVSH 437

RESULT 5

US-10-320-874-16

Sequence 16, Application US/10320874

Publication No. US20030138909A1

GENERAL INFORMATION:

APPLICANT: Koffas, Matheos

APPLICANT: Odom, J. Martin

APPLICANT: No. US20030138909A1ton, Kelley C.

APPLICANT: Ye, Rick

TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN

FILE REFERENCE: CL1619 US NA

CURRENT APPLICATION NUMBER: US/10/320,874

CURRENT FILING DATE: 2002-12-17

PRIOR APPLICATION NUMBER: US/09/934,901

PRIOR FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/229,906

PRIOR FILING DATE: September 1, 2000

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Microsoft Office 97

SEQ ID NO 16

LENGTH: 437

TYPE: PRT

ORGANISM: METHYLOMONAS SP.

US-10-320-874-16

Query Match 100.0%; Score 2284; DB 14; Length 437;

Best Local Similarity 100.0%; Pred. No. 4.8e-207;

Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVTWPHLTADIRFCHEFFLNFTLNKPKKVAIITAGGLAPCLNSAIGSLIRYTE 60

Db 1 DVTWPHLTADIRFCHEFFLNFTLNKPKKVAIITAGGLAPCLNSAIGSLIRYTE 60

Qy 61 IDPSIEIICRYGGYKGLLDGSPYTAEVKKAGVLQRFSGSVIGNSRYKLTNVKDCVXR 120

Db 61 IDPSIEIICRYGGYKGLLDGSPYTAEVKKAGVLQRFSGSVIGNSRYKLTNVKDCVXR 120

Qy 121 GLVKEGEDPQKVAADLVKDGVDIILHTIGDDTNTAAADLAFLANNNGLTIVIGPKTV 180

Db 121 GLVKEGEDPQKVAADLVKDGVDIILHTIGDDTNTAAADLAFLANNNGLTIVIGPKTV 180

Qy 181 DNDVPEIKOSLGAMTAEQAGARYFMNVVAENNANPMLIVHEVMGNCMLTAATQOER 240

Db 181 DNDVPEIKOSLGAMTAEQAGARYFMNVVAENNANPMLIVHEVMGNCMLTAATQOER 240

Qy 241 KLLDRAEWLPELGLTRESYEVAHVFPVEMAIIDLEAKRLREVMDKVDVCNIFVSEGAGV 300

Db 241 KLLDRAEWLPELGLTRESYEVAHVFPVEMAIIDLEAKRLREVMDKVDVCNIFVSEGAGV 300

Qy 301 EAIIVAMQAKGQEVPRDAFGHIKLDVNPCKMFGQFAOMIGAEKTLVOKSGYFAPASAS 360

Db 301 EAIIVAMQAKGQEVPRDAFGHIKLDVNPCKMFGQFAOMIGAEKTLVOKSGYFAPASAS 360

Qy 361 NVDDMLIKSCADLAIECAFRRRESGYIGHEDNGNVLRAIEFPRIKGGKPFNIDTWMFS 420

Db 361 NVDDMLIKSCADLAIECAFRRRESGYIGHEDNGNVLRAIEFPRIKGGKPFNIDTWMFS 420

Qy 421 MUSEIQPKGKVEVSH 437

Db 421 MUSEIQPKGKVEVSH 437

RESULT 6

US-10-363-567-2

Sequence 2, Application US/10363567

Publication No. US20040077068A1

GENERAL INFORMATION:

APPLICANT: E.I. du Pont de Nemours & Company

TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE

FILE REFERENCE: CL1903 PCT

CURRENT APPLICATION NUMBER: US/10/363,567

CURRENT FILING DATE: 2003-09-04

PRIOR APPLICATION NUMBER: US 60/229858

PRIOR FILING DATE: 2000-09-01

PRIOR APPLICATION NUMBER: US 60/229907

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 60

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 437

TYPE: PRT

```

; ORGANISM: Methylobionas 16a
US-10-363-567-2
Query Match      100.0%; Score 2284; DB 15; Length 437;
Best Local Similarity 100.0%; Pred. No. 4,8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWTPYHLTADIRFCHEWFFLNFPYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVTWTPYHLTADIRFCHEWFFLNFPYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEIICVRGGYKGLLGSDSYPTAEVRKKAAGVLOFGSGVIGNSRVKLTNVKDCVCR 120
DB 61 IDPSIEIICVRGGYKGLLGSDSYPTAEVRKKAAGVLOFGSGVIGNSRVKLTNVKDCVCR 120
QY 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAFLARNNGYGLVIGLPTKV 180
DB 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAFLARNNGYGLVIGLPTKV 180
QY 122 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAFLARNNGYGLVIGLPTKV 180
DB 122 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAFLARNNGYGLVIGLPTKV 180
QY 181 DNDVFPKIOSLGAMTAAEQAGARYFMNVVAENNANPRMLIVHEVMGRNCGMLTAAATQOYR 240
DB 181 DNDVFPKIOSLGAMTAAEQAGARYFMNVVAENNANPRMLIVHEVMGRNCGMLTAAATQOYR 240
QY 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNTFVSEBAGV 300
DB 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNTFVSEBAGV 300
QY 301 EAIIVAMQAKQOEVRDPAFGHIKLDVNPCKMFGEQFAQMI GAETLVQKSGYFARASAS 360
DB 301 EAIIVAMQAKQOEVRDPAFGHIKLDVNPCKMFGEQFAQMI GAETLVQKSGYFARASAS 360
QY 361 NVDDMRLLKSCADLAVECAFRESGVI GHDEDNQVLAIEFPRIKGGKPNIDTDWENS 420
DB 361 NVDDMRLLKSCADLAVECAFRESGVI GHDEDNQVLAIEFPRIKGGKPNIDTDWENS 420
QY 421 MLSEIGPKGKVEVSH 437
DB 421 MLSEIGPKGKVEVSH 437

RESULT 7
US-10-321-210-16
; Sequence 16, Application US/10321210
; Publication No. US20040115657A1
; GENERAL INFORMATION:
; APPLICANT: Kofeas, Mattheos
; APPLICANT: Odom, J. Martin
; APPLICANT: Norton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL619 US NA
; CURRENT APPLICATION NUMBER: US/10/321,210
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 437
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-321-210-16
Query Match      100.0%; Score 2284; DB 16; Length 437;
Best Local Similarity 100.0%; Pred. No. 4,8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWTPYHLTADIRFCHEWFFLNFPYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVTWTPYHLTADIRFCHEWFFLNFPYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
```

```

QY 61 IDPSIEIICVRGGYKGLLGSDSYPTAEVRKKAAGVLOFGSGVIGNSRVKLTNVKDCVCR 120
DB 61 IDPSIEIICVRGGYKGLLGSDSYPTAEVRKKAAGVLOFGSGVIGNSRVKLTNVKDCVCR 120
QY 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAFLARNNGYGLVIGLPTKV 180
DB 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAFLARNNGYGLVIGLPTKV 180
QY 122 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAFLARNNGYGLVIGLPTKV 180
DB 122 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAFLARNNGYGLVIGLPTKV 180
QY 181 DNDVFPKIOSLGAMTAAEQAGARYFMNVVAENNANPRMLIVHEVMGRNCGMLTAAATQOYR 240
DB 181 DNDVFPKIOSLGAMTAAEQAGARYFMNVVAENNANPRMLIVHEVMGRNCGMLTAAATQOYR 240
QY 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNTFVSEBAGV 300
DB 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNTFVSEBAGV 300
QY 301 EAIIVAMQAKQOEVRDPAFGHIKLDVNPCKMFGEQFAQMI GAETLVQKSGYFARASAS 360
DB 301 EAIIVAMQAKQOEVRDPAFGHIKLDVNPCKMFGEQFAQMI GAETLVQKSGYFARASAS 360
QY 361 NVDDMRLLKSCADLAVECAFRESGVI GHDEDNQVLAIEFPRIKGGKPNIDTDWENS 420
DB 361 NVDDMRLLKSCADLAVECAFRESGVI GHDEDNQVLAIEFPRIKGGKPNIDTDWENS 420
QY 421 MLSEIGPKGKVEVSH 437
DB 421 MLSEIGPKGKVEVSH 437

RESULT 8
US-10-369-493-14348
; Sequence 14348, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14348
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14348
Query Match      28.9%; Score 659; DB 15; Length 184;
Best Local Similarity 67.9%; Pred. No. 6,9e-54;
Matches 125; Conservative 27; Mismatches 34; Indels 0; Gaps 0;

QY 33 KVAIITAGGLAPCLNSAIGSLIERYTEIDPSIEIICVRGGYKGLLGSDSYPTAEVRK 92
DB 1 QVAMLTAGGLAPCLNSAIGSLIERYSIDIAPEITDIARSGYGVGLGERIITDMKEX 60
QY 93 AGVLOFGSGVIGNSRVKLTNVKDCVCRGLVKGEDPQKVAADQVLDVGDVLIHTIGDD 152
DB 61 AHLHRYGSSPIGNSRVKLTNVKDCVCRGLVKGEDPQKVAADQVLDVGDVLIHTIGDD 120
QY 121 TMTAAADLAFLARNNGYGLVIGLPTKVNDVFPKIOSLGAMTAAEQAGARYFMNVVAENN 212
DB 121 TMTAAADLAFLARNNGYGLVIGLPTKVNDVFPKIOSLGAMTAAEQAGARYFMNVVAENN 212
QY 213 ANPR 216
DB 181 AAPK 184
```

RESULT 9
US-10-369-493-8108
; Sequence 8108, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8108
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Thermobifida fusca
US-10-369-493-8108

Query Match 13.4%; Score 306.5; DB 15; Length 350;
Best Local Similarity 26.5%; Pred. No. 3.8e-20;
Matches 104; Conservative 63; Mismatches 145; Indels 81; Gaps 14;

QY 23 FNFYTLNKKRKAVALTAGGLAPCLNSAIGSLERYEIDPSIITICRGYKGLIGDS 82
DB 4 FKEVTOM---RVGLTGGGDCPGLNAVIRAVRKGIK-EYGEYFVGFRDGMGRPLEGDT 58
QY 83 YPTAERKXKAGVLORFSGSVIGNSRVKLTNVKDCVRGLVKEGEDPQKVAADLVKDV 142
DB 59 MRLDIEVR--GILPR-GGITLSSRTNLMKIBGVER-----VKDMALGV 103
QY 143 DILHTTGDDTNTAAADLAFLLANNYGLTVIGLPTVNDVPEIKOSLGAMTAEGGAR 202
DB 104 DALVAIGEDTLGVARQL-----HDGVNVVGVPKTIDNDLNTDVTFFGDTAVNATE 157
QY 203 YFNNVVAENNAPRMLIVHEVMGNCMLT-----AATQBYRKLDRAEMLPELGITRE 257
DB 158 AIDRLHTTASSHRALTIV-EVMGRHAGMIALHAGMAGA----- 195
QY 258 SYEVHAAVFVPEMADLEAKRLBVDKDCVNIFFVSEGAGVEAIVAEWQAGQEVPRD 317
DB 196 ---NVLLIPERPDIDVVAAYIESRFTNYAPITVVAEGHPRK---EGQLTASGERD 247
QY 318 AFGHIKLDVNPCKMFGGEOPAQMIGAEKTLVQKS---GYPARASASVNDMLIKSCADL 374
DB 248 SFGHVRIGGI-----GQRLAEIEARTCKEARSVNLGHVGRGTSPAFRVLATRLGLH 301
QY 375 AVECAFRRESGVIGHEDNGNV--LRAIEFPRI 405
DB 302 AITAV-----HDKDFGMVALRGTEIYRV 325

RESULT 10
US-10-156-761-13610
; Sequence 13610, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13610
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13610

Query Match 12.8%; Score 294; DB 14; Length 342;
Best Local Similarity 26.5%; Pred. No. 5.7e-19;
Matches 101; Conservative 66; Mismatches 140; Indels 74; Gaps 17;

QY 34 KVALITNAGLAPCLNSAIGSLERYEIDPSIITICRGYKRLIGDSYPTAERKKA 93
DB 2 RVGLTGGGDCPGLNAVIRGVKRVQ-EYGYDFVGFDRDGMGRPLEGDA--VRLDIPAVR 58
QY 94 GVLORFGGSVIGNSRVKLTNVKDCVRGLVKEGEDPQKVAADLVKDVILHTIGGDDT 153
DB 59 GILPR-GGITLSSRTNPLKLDGIR--IKE-----NLAKQEVDAIATIGEDT 105
QY 154 NTAADLAAFLANNYGLTVIGLPTVNDVPEIKOSLGAMTAEGGARF--MNVVAEN 211
DB 106 LGVAARLT-----DEGYFVGVPKTIDNDLSATDVTFFGDTAVGATEAIDRLHTAES 160
QY 212 NANPRMLIVHEVMGRCNGMTAATQBYRKLDRAEMLPELGITRESYEVHAFVPEMAI 271
DB 161 HMR--VLVEVGRHAGMIALHS-----GLAGGA--NVLLIPERQF 197
QY 272 DLEAAKRLBVDKDCVNIFFVSEGAGVEAIVAEWQAGQEVPR-----DAFGHIKLDV 327
DB 198 DVUQCAVYTSRKRASAPRVVAEGA-----MPKQGMPLKDESLDSFGHVRISGV 249
QY 328 NPKMFGGEOPAQMIGAE-KTLVQKSGYPARASASVNDMLIKSCADLAVECAFRRESGV 386
DB 250 --GEMLAKBIEKRTKEARFTV--LGHVGRGTSPAFRVLATRFGLHAIEAV----- 298
QY 387 IGHEDNGNV--LRAIEFPRI 405
DB 299 --RDGDFGMVALRGTDIYRV 317

RESULT 11
US-10-425-114-51700
; Sequence 51700, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 51700
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700837684_FLI.pep
US-10-425-114-51700

Query Match 12.8%; Score 293; DB 15; Length 414;
Best Local Similarity 26.8%; Pred. No. 9.3e-19;
Matches 113; Conservative 61; Mismatches 174; Indels 74; Gaps 15;

QY	34	KVALILNAGGLAPCLNSAIGSLERBERTELDPLEIICVYRGYKGLLDDSYFVTAEVRKKA	93
		: :	
Db	11	QAAIVTCGGGLCPGLNTVLRRELVCGLHHMYGVYKVLGNGYRGFYARNITTLTP--KSV	67
QY	94	GVLAQRFEGSVIGNSRYKLTNNVADCVYKGLVYEGEDPOKVAADQVLKVDVILHTIGDDT	153
		: :	
Db	68	NDIHKKRGVTLGRSR-----GGHDTKKI-VDSIQDRGINQVYIIGDDG	110
QY	154	NTAAADLAAFLARNNYGLTVIGLPTVNDVFPKIQSIGANTAAEQARFPMNVVAENNA	213
		: :	
Db	111	QKASAIIFEEVRRKGLKSVVGIPPTINDIPVIDKSGFDTAAEEQRIINAAHVAES	170
QY	214	NPRMLLVYHEWGRNCMLTAAQAEYRKLDRAEVLPELGTRESYEVAHAFVPEMAIDL	273
		: :	
Db	171	VENGIQVLYKMGRRNSFLA-----MYTTLASRDVDCCLIPESPFYL	211
QY	274	EAEA-----KRLREVMKDVDCVNI FVSEGAQEVAIAEMQAKGQEVPRDAFGHIKIDA	326
		: :	
Db	212	EGCGGLYEVIERKLTKEKGHMV-----IYIAEAGQELVAESVQSMKQ--DASGNKLFQD	264
QY	327	VNNGKVFEGQFQACMIGAEKTL-----VQSKGFAPAASANNVDDMRILKSCADLAVECAFR	381
		: :	
Db	265	V--GLWISQIRDHFAQKTLPTLTLYIDPTMIRAPISNANSDNVYCTLLAQSHAGAM-321	
QY	382	RESGVIGHDED--NGNVLRALIEFPRIKGGKPENIDTD--WFNSMLSEIGOP-----KGGK	432
		: :	
Db	322	--AGTYGYISGLVNGR-QTYIPFYHITERQNHHVITIDMMAR-LISSNQPSFLDAKGDN	377
QY	433	VE 434	
Db	378	EE 379	

```

RESULT 12
US-10-425-114-44512
; Sequence 44512, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313) B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44512
; LENGTH: 525
; TYPE: prt
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700994269_FLI.pep
US-10-425-114-44512

```

[illegible]

Dh 222 QKQASAIIEBVKRRLGKLVSVGICPIPTIONDI.PVIDKSGFPTVAEBQRIINAHAUEAS 281

Qy 214 NRPMLIVHEWGRNCGLTAAETAOEYKLDLDAEMLBELGLTRESYVAHAVPPEMAIDL 273

Dh 282 VENGIGVVLKLMORNSGFLA-----MATTLASRVDCCLLIPESBFYL 322

Qy 274 EABA-----KRLREVMIDKDCVNI.FVSEGAQVEAIVAEQAKGQEPVPRDAFGHILDLA 326

Dh 323 EGGCGI.YEYIEKRLKENHNV---IYIAEBAQGEIVSESVQSMKQ---DASGNKLFQD 375

Qy 327 VNPGRKFGGQFPQMIGAKTIL-----VQSGYIPARAASNVDDMKRLIKSCADLAIVEAFR 381

Dh 376 V--GLMISOKIRIDHRAAQKTLPIYLKIDPTLYMIRALIPSNASDVVYCTLLAQSAHVAGAM- 432

Qy 382 RESGVIGHED--NGNVLRALIEFPRIKGGKFPNIDTD--WENSMUSEIGQP-----KGK 432

Dh 433 --AGYGYISGLVNGR-QTYIPIFYRITERQNHVVITDBMAR--LLSSTNOPSFLAKGDN 488

Qy 433 VE 434

Dh 489 EE 490

```

RESULT 13
US-10-424-599-283989
; Sequence 283989, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 283989
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) -(527)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_98468C.1.pep
US-10-424-599-283989

```

Query Match	Best Local Similarity	12.8% ; Score 293 ; DB 15 ;	Length 527 ;
Matches 113 ;	Conservative 61 ;	Mismatches 174 ;	Indels 74 ; Gaps 15 ;
QY	34	KVAILTAGLAPCLNSAIGSLIERYTEIDPSIEIICRYGGYKGLIGDSTVPTVAERKKA	93
DB	123	QAAIVTCGGCLCPGANTVIRELVCLAHMYGVKVLGINGSGYRGFYARNTITLTP---KSV	179
QY	94	GVLRFGSGVIGNSRVKLTINVKDCVRKGLVKEGEDPOKVAADQVVKDQVDVILHTIGGDDT	153
DB	180	NDIHRGGTVLGTGR-----CGHDTKKL-VDSIQDRINQVYIIIGGGT	222
QY	154	NTAAADLAALFARNNGYLTVIGLPRKYVDNDVFPFKOSIGANTAAEQGARFYMMNVAAENNA	213
DB	223	QKGSAIFEEERKRGKGLKAVSVVGIKTIIDNDIPVIDKSGFPTAAEAORAINAAHVEAES	282
QY	214	NPRALIVHEVNRGCMVLTATAQAEYRKLDRAEMLPGLGTRESYEHAFAVFPEMAIDL	273
DB	283	VENGIQVYKLMGRNSGFIA-----MYATLASRDVDCLIPSPFPYL	323
QY	274	EAEA-----KRLREYMDKVDVCNIPVBSGAGYEALVAENQAQGEVPRDAFGHILKIDA	326
DB	324	EGPGGLVEYIEKRLKENGHNV-----IYAAEGAGELVSESVQSMKQ---DASGKRLFOD	376

Qy 327 VNPGRKFGQFAOMIGAETL-----VOKSGYFARASASVNDMLIKSCADLAECAPR 381
 Db 377 V--GLMISQKIRDHFAAQKTLPTLTKYIDPTMIRALPSNADSNVCTLLAQSAAVHGA-- 433
 Qy 382 RSSGVIGHED--NGNVLAIEFPRIKGGKFPRIIDT--WFSNMLSEIQP-----KGK 432
 Db 434 --AGTYGYSGLVNGR--QTYIPFYRITERONHVITDRMMAR--LSTSTQPSFLDAKGN 489
 Qy 433 VE 434
 Db 490 EE 491

RESULT 14
 US-10-425-114-49830
 / Sequence 49830, Application US/10425114
 / Publication No. US20040034888A1
 / GENERAL INFORMATION:
 / APPLICANT: Liu, Jingdong
 / APPLICANT: Zhou, Yihua
 / APPLICANT: Kovalic, David K.
 / APPLICANT: Screen, Steven E.
 / APPLICANT: Tabaka, Jack E.
 / APPLICANT: Cao, Yongwei
 / TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 / TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 / FILE REFERENCE: 38-21(53313)B
 / CURRENT APPLICATION NUMBER: US/10/425,114
 / CURRENT FILING DATE: 2003-04-28
 / NUMBER OF SEQ ID NOS: 73128
 / SEQ ID NO 49830
 / LENGTH: 527
 / TYPE: PRP
 / ORGANISM: Glycine max
 / FEATURE:
 / OTHER INFORMATION: Clone ID: 700559917_FLI.pcp
 / US-10-425-114-49830

Query Match 12.8%; Score 293; DB 15; Length 527;
 Best Local Similarity 26.8%; Pred. No. 1.3e-18;
 Matches 113; Conservative 61; Mismatches 174; Indels 74; Gaps 15;
 Qy 34 KVALTAGLAPCLNSAIGSLERTYIDPSIEIICRGYKGLLDGDSYPTAVERKKA 93
 Db 124 QAAITVCGGLCPGLMIVRELVGGLHMYGVKKVLGNGYRGFYARNITTLTP---KGV 180
 Qy 94 GVLQRFSGSVIGNSRVKLTNVKDCVKGGLVKGEDPQKVAADQLVKDGVDIHTTIGDDT 153
 Db 181 NDIHKRGCTVLTGSR-----GGHDTKKT-VDSIQRGINQVYIIGDGT 223
 Qy 154 NTPAADLAAFLARNYGLTVIGLPTVDNDVPPIKQSLGAMTPAEGAGYFMNVVAENNA 213
 Db 224 QKGAISLIEPVRKRGKAVSVGIPKTIIDVIDIDSFGDTAVEEAQRIINAAYEAS 283
 Qy 214 NPMRLIVHEVMGNCMLTAATAQERYKLLDRAEWLPELGLTRSEYEVAAVPEMAIDL 273
 Db 284 VENGIGVVLGMGRNSGFI-----WYATLASNDVCCCLIPESPFL 324
 Qy 274 EABA-----KRLREVMKVDVNI-FVSEGAVEAIVAEWQAKGQEVPRDAFGHIKLLDA 326
 Db 325 EEPGGLYEYIERKRLKENGHNV---IYIAEGAGQELVSESVOGMSKQ---DASGNLFPD 377
 Qy 327 VNPGRKFGQFAOMIGAETL-----VOKSGYFARASASVNDMLIKSCADLAECAPR 381
 Db 378 V--GLMISQKIRDHFAAQKTLPTLTKYIDPTMIRALPSNADSNVCTLLAQSAAVHGA-- 434
 Qy 382 RSSGVIGHED--NGNVLAIEFPRIKGGKFPRIIDT--WFSNMLSEIQP-----KGK 432
 Db 435 --AGTYGYSGLVNGR--QTYIPFYRITERONHVITDRMMAR--LSTSTQPSFLDAKGN 490
 Qy 433 VE 434

Db 491 EE 492
 RESULT 15
 US-10-156-761-14652
 / Sequence 14652, Application US/10156761
 / Publication No. US20030119018A1
 / GENERAL INFORMATION:
 / APPLICANT: OMURA, SATOSHI
 / APPLICANT: IKEDA, HARUO
 / APPLICANT: ISHIKAWA, JUN
 / APPLICANT: HORIKAWA, HIROSHI
 / APPLICANT: SHIBA, TADAYOSHI
 / APPLICANT: SAKAKI, YOSHIYUKI
 / APPLICANT: HATTORI, MASAHIRA
 / TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 / FILE REFERENCE: 249-262
 / CURRENT APPLICATION NUMBER: US/10/156,761
 / CURRENT FILING DATE: 2002-05-29
 / PRIOR APPLICATION NUMBER: JP 2001-204089
 / PRIOR FILING DATE: 2001-05-30
 / PRIOR APPLICATION NUMBER: JP 2001-272697
 / PRIOR FILING DATE: 2001-08-02
 / NUMBER OF SEQ ID NOS: 15109
 / SEQ ID NO 14652
 / LENGTH: 341
 / TYPE: PRP
 / ORGANISM: Streptomyces avermitilis
 / US-10-156-761-14652

Query Match 12.5%; Score 284.5; DB 14; Length 341;
 Best Local Similarity 27.5%; Pred. No. 4.5e-18;
 Matches 97; Conservative 62; Mismatches 143; Indels 51; Gaps 11;
 Qy 34 KVALTAGLAPCLNSAIGSLERTYIDPSIEIICRGYKGLLDGDSYPTAVERKKA 93
 Db 2 RIGVLTSGDCGCLNAVIRSVHR--AVVDHGDVIGFRDGMKGLLECDY--LKLDDAVS 58
 Qy 94 GVLQRFSGSVIGNSRVKLTNVKDCVKGGLVKGEDPQKVAADQLVKDGVDIHTTIGDDT 153
 Db 59 GILAR--GGTITLSSSRVQPAHLDDVER-----ARGHVAELGIDAIIPIGGBGT 105
 Qy 154 NTPAADLAAFLARNYGLTVIGLPTVDNDVPPIKQSLGAMTPAEGAGYFMNVVAENNA 213
 Db 106 LKAARLL-----SPAGLPIVGVPTKTIIDNDIAVTDTFGDTAVGATBALDLKTTAAS 159
 Qy 214 NPMRLIVHEVMGNCMLTAATAQERYKLLDRAEWLPELGLTRSEYEVAAVPEMAIDL 273
 Db 160 HORVLTIV--EVMGRHGTMLHSGM-----AAGAHAIIVPERPFDI 198
 Qy 274 EABA--KRLREVMKVDVNI-FV--SEGAVEAIVAEWQAKGQEVPRDAFGHIKLLDAVNPGRK 332
 Db 199 EELAAKVGRRFPAAGKRAIIVAABGAKKPPAGSMDF---DEGKDYVGHERRPAGT--AAQ 252
 Qy 333 FGEQFAOMIGAETLVOKSGYFARASASVNDMLIKSCADLAECAPRREG 385
 Db 253 LSLIEERLKGKARPV--ILGHVQGGTPTAAYRVLATRFGHVAEVAHVHGEFG 304

Search completed: February 24, 2005, 14:34:59
 Job time : 135 secs

THIS PAGE BLANK (USP)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using ew model

Run on: February 24, 2005, 14:01:14 ; Search time 182 Seconds

(without alignments)
1229.553 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTMPHYLTADIRFCWFF.....FNSMLSEIGQPKGKGVESH 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1531.5	67.1	453	Q7UES2	Q7UES2 mastodirell
2	1453	63.6	410	Q9NGP6	Q9NGP6 mastigamob
3	1432	62.7	399	Q6A8S9	Q6A8S9 propionbac
4	1314	57.5	403	PPF PROFR	P29495 p pyrophosp
5	1302	57.0	404	Q92NK9	Q92NK9 rhizobium m
6	1292.5	56.6	479	Q8UDL0	Q8UDL0 agrobacteri
7	1292.5	56.6	514	Q7C6V0	Q7C6V0 agrobacteri
8	311	13.6	342	1 K6P1 STRCO	Q08333 streptomyc
9	305	13.4	346	1 PFP DICTH	Q9KH71 d pyrophosp
10	298.5	13.1	341	1 PFP AMYME	Q59126 a pyrophosp
11	294	12.9	342	2 Q82NH4	Q82NH4 streptomyc
12	284.5	12.5	341	2 Q82E03	Q82E03 streptomyc
13	282	12.3	488	2 Q9FYU5	Q9FYU5 arabisdopsi
14	281.5	12.3	341	1 K6P3 STRCO	Q9FJUS arabisdopsi
15	281.5	12.3	485	2 Q9CSJ7	Q9CSJ7 streptomyc
16	279.5	12.2	358	2 Q6AG22	Q6AG22 leifsonia x
17	278.5	12.2	542	2 Q6LAN6	Q6LAN6 oryza sativ
18	278	12.2	366	1 K6P2 CLOPE	Q8X157 clostridium
19	275	12.0	447	2 Q660E6	Q660E6 borrelia ga
20	274.5	12.0	341	1 PFP AMYMD	Q9AGC0 a pyrophosp
21	271.5	11.9	341	1 K6P2 STRCO	Q911I8 streptomyc
22	270	11.8	448	2 Q516E9	Q516E9 borrelia bu
23	269.5	11.8	341	2 Q82UD3	Q82UD3 streptomyc
24	266.5	11.7	319	2 Q74BH3	Q74BH3 geobacter s
25	266	11.6	489	2 Q94AA4	Q94AA4 arabisdopsi
26	263.5	11.5	436	2 Q276S1	Q276S1 entamoeba h
27	263.5	11.5	586	2 Q94UJ2	Q94UJ2 oryza sativ
28	260.5	11.4	775	2 Q9UG72	Q9UG72 oryza sativ
29	260	11.4	382	2 Q248I2	Q248I2 entamoeba h
30	259.5	11.4	444	2 Q72AD6	Q72AD6 desulfovibr
31	257.5	11.3	500	2 Q9ST07	Q9ST07 arabisdopsi

32	256.5	11.2	567	2 Q65X97	Q65X97 oryza sativ
33	256.5	11.2	991	1 Q6FTX6	Q6FTX6 candida gla
34	255.5	11.2	319	1 K6P2 CLOTE	Q89022 clostridium
35	253.5	11.1	987	2 Q758H0	Q758H0 ashbya gose
36	253	11.1	319	1 K6P2 THEM4	Q9WY52 thermocoga
37	253	11.1	473	2 Q817I4	Q817I4 arabisdopsi
38	252	11.0	459	2 Q8VU09	Q8VU09 amycolatops
39	252	11.0	487	2 Q15648	Q15648 trypanosoma
40	251	11.0	473	2 Q9M0F9	Q9M0F9 arabisdopsi
41	250	10.9	426	2 Q61068	Q61068 trichomonas
42	248.5	10.9	437	2 Q27705	Q27705 naegleria f
43	246.5	10.8	987	1 K6P1 YEAST	P16861 saccharomyc
44	245.5	10.7	992	1 K6P1 KUTU4	Q03215 Kluyveromyc
45	245	10.7	324	2 Q8GNC1	Q8GNC1 haemophilus

ALIGNMENTS

RESULT 1

ID	Q7UES2	PRELIMINARY;	PRT;	453 AA.
AC	Q7UES2;			
DT	01-OCT-2003 (TRMBLrel. 25, Created)			
DT	01-OCT-2003 (TRMBLrel. 25, Last sequence update)			
DT	01-MAR-2004 (TRMBLrel. 26, Last annotation update)			
DE	PPI-phosphofructokinase (EC 2.7.1.90).			
GN	Name=Pfk; OrderedLocusNames=RB10591;			
OS	Rhodospirillum rubrum.			
OC	Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;			
OC	Planctomycetaceae; Pirellula.			
OX	NCBI_TaxID=117;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=1;			
RX	MEDLINE=22735913; PubMed12835416; DOI=10.1073/pnas.1431443100;			
RA	Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,			
RA	Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,			
RA	Schlesner H., Amann R., Reinhardt R.;			
RT	"Complete genome sequence of the marine planctomycete Pirellula sp.			
RT	strain 1."			
RL	Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).			
DR	EMBL; BX294151; CAD78962.1; -.			
DR	GO; GO:0005945; C:6-phosphofructokinase complex; IEA.			
DR	GO; GO:0003872; F:6-phosphofructokinase activity; IEA.			
DR	GO; GO:0047334; F:diphosphate-fructose-6-phosphate 1-phosphot. . .; IEA.			
DR	GO; GO:0016301; F:kinase activity; IEA.			
DR	GO; GO:0016740; F:transferase activity; IEA.			
DR	GO; GO:0006096; P:glycolysis; IEA.			
DR	InterPro; IPR000023; Pfkfructinase.			
DR	Pfam; PF00365; Pfk; 1.			
DR	Prodom; PD000707; PpfFructinase; 1.			
KW	Complete proteome; Kinase; Transferase.			
SEQ	SEQUENCE 453 AA; 49268 MW; 9D7D76C8CC909DAB CRC64;			

Query Match 67.1%; Score 1531.5; DB 2; Length 453;
Best Local Similarity 65.9%; Pred. No. 3.5e-88;
Matches 286; Conservative 61; Mismatches 78; Indels 9; Gaps 1;

QY	4	TWVPHLTADIRFCWFFLNENFYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTEIDP	63
DB	29	SWKIHAFASLESEPIIMSI-----KRVGILTLAGLAPCLSSAIGALIEAYEQAP	79
QY	64	STETTCYRGYGGLLGDSYPYTAAYRKAACVLRFGSGVINSRVKTNNVDCVRRGIV	123
DB	80	ETETTCYRSYGGYLLGDSFVVDASRKNAALIHGGSPGNSRVKLTNNVADCVRRGIV	139
QY	124	KGSDPQKVAADQVLDVGDVLTATIGGDDTNTAAADLAFLARNNYGLTVIGLPTVDND	183
DB	140	SGGDPQLQVAARLQSDVDVLTATIGGDDTNTAADLAFLAKHEQLTVVGPKRIDND	199
QY	184	VPTIKOSLGAWTAAGARFYEMNVVVAENNANFRMLIVHEVMGRNGCWLTAATAOEYRKLL	243

DB 200 VIFPIQSIGANTAAEGAKFFENVYGERHNPRLIIEHWGRNCGMLTAATATYRRL 259
QY 244 DRAEWLPGLGTRRESYEVHNAVFPVEMAIIDLEAKRLBEVMDKVCNIVFVSEGAQVBI 303
DB 260 QULNLPFEMGJLSQBRDVGAVFPFEMSFDLEBAKRLAIIMEIDCVNIFISEGAGVDTI 319
QY 304 VAEMAKQOEVPDRDFAFGHKLDAVNPGRKFGQFQPMQMGAEKTLVQKSGYFARASAVD 363
DB 320 VKEMSRGESVPKDFGHKLDAVNPGRKFGQFQPMQMGAEKTLVQKSGYFARASAVD 379
QY 364 DMRLIKSCADLAVECAFRRRESGVIGHDEDNVLAIEFPRIKGGKFPNIDTFNSMLS 423
DB 380 DIALIGRCAKKAVECAQMGQIGSVGEDEDDNMLAIEFRIAGKFPDINWVFGDLS 439
QY 424 EIGQPKGKVEVSH 437
DB 440 QMGQPKGEVLTESH 453

RESULT 2

Q9NGP6 PRELIMINARY; PRT; 410 AA.
AC Q9NGP6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE PPI-phosphofructokinase (EC 2.7.1.11).
GN Name=Pp1pfk;
OS Maestigamoeba balamuthi (Phreatamoeba balamuthi).
OC Eukaryota; Pelobiontida; Maestigameobidae; Maestigameoba.
NCBI_Taxid=108607;
RX MEDLINE=21528850; PubMed=11673446;
DOI=10.1128/JB.183.22.6714-6716.2001;
RA Muller M., Lee J.A., Gordon P., Gaasterland T., Sensen C.W.:
"Presence of prokaryotic and eukaryotic species in all subgroups of
RT the Pp1(1)-dependent group II phosphofructokinase protein family.";
RL J. Bacteriol. 183:6714-6716(2001).
DR EMBL; AF246209; AAF70463.1; -
DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
DR GO; GO:0003872; F:6-phosphofructokinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000023; Pfruckinase.
DR InterPro; IPR011405; PPI-PFK_SMC01852.
DR Pfam; PF00365; PFK; 1.
DR PIRSF; PIRSF036484; PPI-PFK_SMC01852; 1.
DR PRINTS; PR00476; PHERCTKINASE.
DR KMW Kinase; Transferase.
SQ SEQUENCE 410 AA; 44200 MW; CD362D1D2D41A0D0 CRC64;

Query Match 63.6%; Score 1453; DB 2; Length 410;
Best Local Similarity 67.2%; Pred. No. 2.7e-83;
Matches 272; Conservative 55; Mismatches 78; Indels 0; Gaps 0;

QY 33 KKVAILITAGGLAPCLNSAIGSLIBRYTEIDPSIEIICRYGKYGLLDGSPVTAVERK 92
DB 3 KTVALLITAGGLAPCLNSAIGSLIBRYTEIDPSIEIICRYGKYGLLDGSPVTAVERK 62
QY 93 AGVILQRFSGSVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQIVKQGVILHTIGDD 152
DB 63 AAVLHTVGGSCIGNSRVKMANVADCVKGLVKEGDPQVADQILIKQGVILHTIGDD 122
QY 153 TMTAAADILAAFLAANNVGLTVIGLPTVNDVFPKQSLGAWTAEGQARFMNVVAEN 212
DB 123 TMTAAADILAAFLAANNVGLTVIGLPTVNDVFPKQSLGAWTAEGQARFMNVVAEN 182
QY 213 ANPRLIYHEVWGRNCGMLTAATAOEYRKLDRAEWLPGLGTRRESYEVHNAVFPVEMAI 272
DB 183 ANPRLIYHEVWGRNCGMLTAATAOEYRKLDRAEWLPGLGTRRESYEVHNAVFPVEMAI 242

QY 273 LEBAKRLREVMKVDCVNIIVFVSEGAQVEALVAEMQAGQEVPRDAFGHKLDAVNPGRK 332
DB 243 LKAAARLRARVAMERVCNIFLSEGAQVINDIVAEKTAAGTVPRDPFGHVKIDLINPAW 302
QY 333 FGEQFAQMGAEKTLVQKSGYFARASAVNDMLIKSCADLAVECAFRRRESGVIGHDE 392
DB 303 FGKQFGGVAGDKTLVQKSGYFARASAVNADMLIKSCADLAVECAFRRRESGVIGHDE 362
QY 393 NGNVLAIEFPRIKGGKFPNIDTFNSMLSEIQPKGKVEVSH 437
DB 363 RNVGLRAIEFPRVYKAKAFNIDHPWFTLHNEIQPKAKVSVAH 407

RESULT 3

Q6A8S9 PRELIMINARY; PRT; 399 AA.
AC Q6A8S9;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Pyrophosphate--fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.90).
GN OrderedLocNames=PPA1090;
OS Propionibacterium acnes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.
NCBI_Taxid=1747;
RX SEQUENCE FROM N.A.
RC STRAIN=KPA173202 / DSM 16379;
PubMed=15286373; DOI=10.1126/science.1100330;
RA Brueggemann H., Henne A., Hostler F., Liesegang H., Wielez A.,
RA Stittmattler A., Hujer S., Duerre P., Gottschalk G.:
"The complete genome sequence of Propionibacterium acnes, a commensal
RT of human skin.";
RL Science 305:671-673(2004).
DR EMBL; AE017283; AAT82837.1; -
DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
DR GO; GO:0003872; F:6-phosphofructokinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000023; Pfruckinase.
DR Pfam; PF00365; PFK; 1.
DR PRINTS; PR00476; PHERCTKINASE.
DR ProDom; PD000707; Pfruckinase; 1.
DR KMW Complete proteome; Transferase.
SQ SEQUENCE 399 AA; 42344 MW; A0A845D7FAB0C1CD CRC64;

Query Match 62.7%; Score 1432; DB 2; Length 399;
Best Local Similarity 68.4%; Pred. No. 5.4e-82;
Matches 270; Conservative 54; Mismatches 71; Indels 0; Gaps 0;

QY 33 KKVAILITAGGLAPCLNSAIGSLIBRYTEIDPSIEIICRYGKYGLLDGSPVTAVERK 92
DB 4 KKVALLITAGGLAPCLNSTAIGSLIQRYTEVAPBEVILAIKHGVEGLKDDFLEVTDTYK 63
QY 93 AGVILQRFSGSVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQIVKQGVILHTIGDD 152
DB 64 AEILKRFSGSPVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQIVKQGVILHTIGDD 123
QY 153 TMTAAADILAAFLAANNVGLTVIGLPTVNDVFPKQSLGAWTAEGQARFMNVVAEN 212
DB 124 TMTAAADILAAFLAANNVGLTVIGLPTVNDVFPKQSLGAWTAEGQARFMNVVAEN 183
QY 213 ANPRLIYHEVWGRNCGMLTAATAOEYRKLDRAEWLPGLGTRRESYEVHNAVFPVEMAI 272
DB 184 SGRMLIYHEVWGRNCGMLTAATAOEYRKLDRAEWLPGLGTRRESYEVHNAVFPVEMAI 243
QY 273 LEBAKRLREVMKVDCVNIIVFVSEGAQVEALVAEMQAGQEVPRDAFGHKLDAVNPGRK 332
DB 244 LEBAARLRARVAMERVCNIFLSEGAQVINDIVAEKTAAGTVPRDPFGHVKIDLINPAW 303
QY 333 FGEQFAQMGAEKTLVQKSGYFARASAVNDMLIKSCADLAVECAFRRRESGVIGHDE 392

```

Db      304 FGEQFDKJLGAEEVWQKSGYFSRSASASNEADLELLGRCTDLAVDCAAGKTGVIQODEB 363
Qy      393 NGNVLRALIEFPRIKGGKPFNIDTDFNSMLSEIQ 427
Db      364 NGDITLNIAPDRIKGGKPFDTTQPPWTATMLSEIQ 398

RESULT 4
PEP_PROPR
ID_PEP_PROPR      STANDARD:      PRT:      403 AA.
P29195;

DT      01-APR-1993 (rel. 25, Created)
DT      01-APR-1993 (rel. 25, Last sequence update)
DT      25-OCT-2004 (rel. 45, Last annotation update)
DE      Pyrophosphatase--fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.90)
DE      (6-phosphofructokinase, pyrophosphate dependent) (Pyrophosphatase-
DE      dependent 6-phosphofructose-1-kinase) (PPI-dependent
DE      phosphofructokinase) (PPI-PFK).
GN      Name=ppf; Synonyms=pfk;
OS      Propionibacterium freudenreichii shermanii.
OC      Bacteria; Actinobacteria; Actinobacteriales;
OC      Propionibacteriaceae; Propionibacterium.
OX      NCBI_Taxid=1752;

[1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX      MEDLINE=91358443; PubMed=1653240;
RA      Lador U.S., Gollupudl L., Tripathi R.L., Latschew S.P., Kemp R.G.;
RT      "Cloning, sequencing, and expression of pyrophosphate-dependent
RT      phosphofructokinase from Propionibacterium freudenreichii.";
RN      J. Biol. Chem. 266:16550-16555(1991).
[2]
IDENTIFICATION OF CRITICAL LYSYL RESIDUES.
RX      MEDLINE=92273593; PubMed=1317210;
RA      Green P.C., Latschew S.P., Lador U.S., Kemp R.G.;
RT      "Identification of critical lysyl residues in the pyrophosphate-
RT      dependent phosphofructo-1-kinase of Propionibacterium
RT      freudenreichii.";
RN      Biochemistry 31:4815-4821(1992).
-1- CATALYTIC ACTIVITY: Diphosphate + D-fructose 6-phosphate =
-1- phosphate + D-fructose 1,6-bisphosphate.
-1- ENZYME REGULATION: Non-allosteric.
-1- PATHWAY: Key control step of glycolysis.
-1- SUBUNIT: Homodimer.
-1- SIMILARITY: Belongs to the phosphofructokinase family. PPF
subfamily.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement. (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M67447; AAA25675.1; -
DR      PIR; A41169; A41169.
DR      InterPro; IPR000023; PpfFructKinase.
DR      Pfam; PF00365; PFK_1; PPI-PFK_SMC01852; 1.
DR      PRINTS; PIRSF036484; PPI-PFK_SMC01852; 1.
DR      PIRSF; PIRSF036484; PPI-PFK_SMC01852; 1.
DR      ProDom; PD000707; PpfFructKinase; 1.
KW      ATP-binding; Direct protein sequencing; Kinase; Magnesium;
KW      Transference.
FT      INIT MET      0
FT      NP_BIND      21      25      ATP (By similarity).
FT      NP_BIND      120     125      Pyrophosphate (potential).
FT      NP_BIND      178     182      ATP (By similarity).
FT      NP_BIND      195     231      ATP (By similarity).
FT      ACT_SITE      150     150      Proton acceptor (By similarity).
FT      BINDING      325     325      Substrate (By similarity).
SQ      SEQUENCE      403 AA; 43114 MW; 1DEB3272B4A2B6E6 CRC64;

```

```

Query Match      57.5%; Score 1314; DB 1; Length 403;
Best Local Similarity 62.6%; Pred. No. 1,4e-74;
Matches 248; Conservative 60; Mismatches 88; Indels 0; Gaps 0;

Qy      33 KXVALITAGGAPCUNSAIGSLIERTEIDPSLEIICYGKYGKLLGDSYPTAEVRKK 92
Db      2 KXVALITAGGAPCUNSAIGSLIERTEIDPSLEIICYGKYGKLLGDSYPTAEVRKK 61

Qy      93 AGVLRFGSGVIGNSRVLKTNVDCVKGRLVKEGEDPQKVAADQVLKQDVIHTIGDD 152
Db      62 YRLRFSFGSGPIGNSRVLKTNVDCVKGRLVKEGEDPQKVAADQVLKQDVIHTIGDD 121

Qy      153 TWTAAADLAFLARNNGYLVITGLPKTVNDVFPFKQSIGAWTAAGARFYFNVAENN 212
Db      122 TWTAAADLAFLARNNGYLVITGLPKTVNDVFPFKQSIGAWTAAGARFYFNVAENN 181

Qy      213 ANPRMLIVHWNGRNGWLTAAAGCYKLLDRAEKLPELGLTRESYEVHNVPEMAID 272
Db      182 AAPRELIIHEIRNGRNGVLAETSRRYVAMLAQQLPAGLDGRGMDIHALYVPEATID 241

Qy      273 LBAEAKRLREWMDKVDGVNI FVSEGAGVEAIVAEWQAKQGEVPRDAFGHIKLDVNPCKY 332
Db      242 LBAEAKRLREWMDKVDGVNI FVSEGAGVEAIVAEWQAKQGEVPRDAFGHIKLDVNPCKY 301

Qy      333 FGEQFAQMIAGKTLVQKSGYFSRSASASNVDMRLIKSCADLAIVECAFRESGVIGHDE 392
Db      302 FAKQFAERIGAGKTWQKSGYFSRSASNAQDLLELAATATATVADALAGTPTGVGQDEB 361

Qy      393 NGNVLRALIEFPRIKGGKPFNIDTDFNSMLSEIQ 428
Db      362 AGDKLSVIDFKRIAGHKPFDITLDTWYQLARIQGP 397

RESULT 5
Q92NK9
ID_Q92NK9      PRELIMINARY;      PRT:      404 AA.
Q92NK9;
AC      Q92NK9;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      PROBABLE PYROPHOSPHATE--FRUCTOSE 6-PHOSPHATE 1-PHOSPHOTRANSFERASE
DE      PROTEIN (EC 2.7.1.90).
GN      ORFName=SMC01852;
OS      Rhizobium meliloti (Sinorhizobium meliloti).
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX      NCBI_Taxid=382;

[1]
SEQUENCE FROM N.A.
RX      STRAIN=1021;
RX      MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
RA      Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Anpe F., Batut J.,
RA      Boissard P., Becker A., Boutry M., Cadieu E., Dreaño S., Gloux S.,
RA      Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA      Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA      Renard C., Thebaud P., Vandendol M., Weidner S., Galibert F.;
RT      "Analysis of the chromosome sequence of the legume symbiont
RT      Sinorhizobium meliloti strain 1021.";
RN      Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
RL      EMBL; AL591789; CAC46769.1; -
DR      GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
DR      GO; GO:0003872; P:6-phosphofructokinase activity; IEA.
DR      GO; GO:0047334; F:diphosphate-fructose-6-phosphate 1-phosphot. . .; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      CO; GO:0006096; P:glycolysis; IEA.
DR      InterPro; IPR002110; ANK.
DR      InterPro; IPR000023; PpfFructKinase.
DR      InterPro; IPR011405; PPI-PFK_SMC01852.
DR      Pfam; PF00365; PFK_1.
DR      PIRSF; PIRSF036484; PPI-PFK_SMC01852; 1.
DR      PRINTS; PRO1415; ANKYRIN.
DR      PRINTS; PRO0476; PPFCTKINASE.
DR      ProDom; PD000707; PpfFructKinase; 1.

```

KW Complete proteome; Transferase.
 SQ SEQUENCE 404 AA; 43712 MW; 1E6E5CD966A13D5 CRC64;
 Query Match 57.0%; Score 1302; DB 2; Length 404;
 Best Local Similarity 62.3%; Pred. No. 8.1e-74;
 Matches 246; Conservative 59; Mismatches 90; Indels 0; Gaps 0;

QY 33 KKVAILTGGIAPCINSAIGSLIERYTEIDPSIEITICRGYKGLLGDSPVTAEVK 92
 DB 4 KKVAILTGGIAPCINSAIGSLIERYTEIDPSIEITICRGYKGLLGDSPVTAEVK 63
 QY 93 AGVLRFSGSVIGNSRVKLTNNKDCVKRGLVKEGSDPOKVAADOLVKQVDILHTIGDD 152
 DB 64 AHVLRHSGSPVIGNSRVKLTNNKDCVKRGLVKEGSDPOKVAADOLVKQVDILHTIGDD 123
 QY 153 TTTAAADLAFLARNNGYLVIGLPTVDNDVFPFKQSLGAWTAAEGGARVFNVAENN 212
 DB 124 TTTAAADLAFLARNNGYLVIGLPTVDNDVFPFKQSLGAWTAAEGGARVFNVAENN 183
 QY 213 ANPRMLTYHEVWGRNCMLTAAIDBYRKLDRAMLEBELGTRSEYHAFVPEMAID 272
 DB 184 ANPRMLTYHEVWGRNCMLTAAIDBYRKLDRAMLEBELGTRSEYHAFVPEMAID 243
 QY 273 LEAEAKRLREVNDKVDVNIFFVSEAGVYEAIVAEQAKGOEVRDPAFGHIKLDVNPCKM 332
 DB 244 LEAEAKRLREVNDKVDVNIFFVSEAGVYEAIVAEQAKGOEVRDPAFGHIKLDVNPCKM 303
 QY 333 FGEQFQMIQAEKTLVQKSGYFAPASASVNDMDRLIKSCADLAVCAFRRESVIGHED 392
 DB 304 FSKQFALLIGARSNVQKSGYFAPASASVNDMDRLIKSCADLAVCAFRRESVIGHED 363
 QY 393 NGNVLRATEFPRIKGGKPPNITDTPFNMSLSEIGQ 427
 DB 364 OGGRLLTIEFPRIKGGKPPNITDTPFNMSLSEIGQ 398

RESULT 6
 Q8UDL0 PRELIMINARY; PRT; 479 AA.
 ID Q8UDL0
 AC Q8UDL0;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Pyrophosphate--fructose-6-phosphate 1-phosphotransferase.
 GN Name=pfP; OrderedLocNames=Atu2115;
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 NX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Dupont;
 RX MEDLINE=21608550; Pubmed=11743193; DOI=10.1126/science.1066804;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Chou Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. St.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutyavyn T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Seemphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Kespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chundley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Neeser E.W.,
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58.";
 RL Science 294:2317-2323 (2001).
 DR EMBL; AE009161; AAL4105.1; -.
 DR PIR; AC2836; AC2836.
 DR PIR; F97613; F97613.
 DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
 DR GO; GO:0003872; F:6-phosphofructokinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.

DR Pfam: PF00365; PRK: 1.
 DR PRINTS: PR00476; PHPRCKINASE.
 DR Prodom: PD000707; Pfruckinase; 1.
 SQ SEQUENCE 479 AA; 52036 MW; 86934EBF948A2C56 CRC64;
 Query Match 56.6%; Score 1292.5; DB 2; Length 479;
 Best Local Similarity 57.9%; Pred. No. 3.9e-73;
 Matches 246; Conservative 71; Mismatches 101; Indels 7; Gaps 2;

QY 8 HLTAADI-----RFGWFPLNFNYTLNNKPKVAITLTAAGLAPCINSAIGSLIERYTEIDP 63
 DB 54 HLHAANVSIQCTCLMVAK---RRAMAKQKVAULTAGLAPCINSAIGSLIERYTEIDP 110
 QY 64 SIEITICRGYKGLLGDSPVTAEVKKAQVLFQSGSVIGNSRVKLTNNKDCVKRGLV 123
 DB 111 EIDIIAYRSGYQGVILGRIETITDMREKAHLIRYGGSPVIGNSRVKLTNNKDCVKRGLV 170
 QY 124 KEGSDPOKVAADOLVKQVDILHTIGDDTNTAAADLAFLARNNGYLVIGLPTVDND 183
 DB 171 KEGSDPFLVAERLAADGITLHTIGDDTNTAAADLAFLARNNGYLVIGLPTVDND 230
 QY 184 VFPIKQSLGAWTAAEGGARVFNVAENNANPRMLTYHEVWGRNCMLTAAIDBYRKL 243
 DB 231 VFPIKQSLGAWTAAEGGARVFNVAENNANPRMLTYHEVWGRNCMLTAAIDBYRKL 290
 QY 244 DRAEMLEBELGTRSEYHAFVPEMAIDLEAEAKRLREVNDKVDVNIFFVSEAGVYEA 303
 DB 291 RGNVYVEELMNTQMKINDGITLPEMAFIDAEAEKRLREVNDKVDVNIFFVSEAGVYEA 350
 QY 304 VAEMQAKGOEVRDPAFGHIKLDVNPCKMFGEOFQMIQAEKTLVQKSGYFAPASASVND 363
 DB 351 VAERBAEAVKRDPAFGHIKLDVNPCKMFGEOFQMIQAEKTLVQKSGYFAPASASVND 410
 QY 364 DMLRIKSCADLAVCAFRRESVIGHEDNGNVLRATEFPRIKGGKPPNITDTPFNMSLSEIGQ 423
 DB 411 DMLRIKSCADLAVCAFRRESVIGHEDNGNVLRATEFPRIKGGKPPNITDTPFNMSLSEIGQ 470

RESULT 7
 Q7CXV0 PRELIMINARY; PRT; 514 AA.
 ID Q7CXV0
 AC Q7CXV0;
 DT 05-JUL-2004 (Tremblrel. 27, Created)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE AGR_C_3836P.
 GN OrderedLocNames=AGR_C_3836;
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 NX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Cereon;
 RX MEDLINE=21608551; Pubmed=11743194; DOI=10.1126/science.1066803;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
 RA Houtrel K., Gordon J., Vaubin M., Iatchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.,
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328 (2001).
 DR EMBL; AE008127; AAK87863.1; -.
 DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
 DR GO; GO:0003872; F:6-phosphofructokinase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.

```

DR InterPro: IPR000023; Pfructinase.
DR InterPro: IPR011405; PPI-PFK_SMC01852.
DR Pfam: PF00365; PFK, 1.
DR PIRSF: PIRSF03484; PPI-PFK_SMC01852; 1.
DR PRINTS: PRO0476; PHFCKTKINASE.
DR Prodom: PD000707; Pfructinase; 1.
DR SEQUENCE 514 AA; 55840 MW; B92DF77E39F0B0AD CRC64;

Query Match 56.6%; Score 1292.5; DB 2; Length 514;
Best Local Similarity 57.9%; Pred. No. 4.3e-73;
Matches 246; Conservative 71; Mismatches 101; Indels 7; Gaps 2;

QY 8 HLTADI-----RFGHMFLLNPNFTYTLNNKPKCAITLTAAGLAPCLNSAIGSLIRYEIDP 63
DB HHAAVAASICQICLWYAK---RRAMMAKQKVMLTAGGLAPCLNSAVGLIERYDIAP 145
QY 64 STEITCYRGYKGLLGDSPVTAERKKGAVQGRGSGVIGNSRKLTNVDCVKGIV 123
DB 146 EIDILAVRSGYQGVLLGERIEITKDRKXHLHRRGSGPIGNSRYKLTNAADCAKRGIV 205
QY 124 KEGEDPQKVAADQLVKDGVDILHTIGDDPTNTAAADLAFLANNYGLTVIGLPTVND 183
DB 206 KEGDNLRLRAERLEADGITLIHTIGDDPTNTAAADLAFLANGDILTVGLPKTVND 265
QY 184 VPIKQSLGAMTAAEGGARYFMNVVAENNAEMRLIVHEVMGNCMLTAATAGBYRKLL 243
DB 266 VPIKQSLGAMTAAEGGARYFMNVVAENNAEMRLIVHEVMGNCMLTAATAGBYRKLL 325
QY 244 DDAEMLPGLGTRRESYEVAVFPEMAIDLEAFKRLRYMDVDCVNI FVSGAGVEAI 303
DB 326 RNDVYEGMLMNTQMKNIDIGIYLPENAFDIEAERLKEVMDGQVYTLFVSGAGLDAI 385
QY 304 VAEWQAGQEVPPDAFGHILKLDVNPKNFGEQFAOMIGEKTLVOKSGYFARASAND 363
DB 386 VAREKAGAVKRDAGHKIDINVGCMFOKOPAGLIGERMSVOKSGYFARASAND 445
QY 364 DMRLIKSCADLAVECAFRRBSGVIGHEDNGVNLAIIEFPRIKGRPFNIDTWFMSLS 423
DB 446 DRLIGQWDLAVESALNKVSGVTGHEDQNGKLTIFPRIKGKHFDLSAKMFIEVME 505
QY 424 ETGQIP 428
DB 506 HWGQP 510

RESULT 8
KEPI_STRCO STANDARD; PRT; 342 AA.
ID K6PI_STRCO
AC 008333;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE 6-phosphofructokinase 1 (EC 2.7.1.11) (Phosphofructokinase 1)
GN ORPName=SC6E10.13c;
GN ORPName=SC6E10.13c;
OS ORPName=SC6E10.13c;
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces coelicolor.
OC Streptomyces; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1902;
RN (1)
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=A3(2) / 1109;
RX MEDLINE=97208211; PubMed=9055413;
RA Alvee A.M.C.R., Eversink G.V.W., Bibb M.J., Dijkhuizen L.;
RT Identification of ATP-dependent phosphofructokinase as a regulatory
RT step in the glycolytic pathway of the actinomycete Streptomyces
RT coelicolor A3(2).
RL Appl. Environ. Microbiol. 63:956-961(1997).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;

```

```

RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thompson N.R., James K.D., Harris D.E., Quail M.A., Kieseer H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieseer T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinovitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)".
RL Nature 417:141-147(2002).
CC -1- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
CC fructose 1,6-bisphosphate.
CC -1- ENZYME REGULATION: Allosterically inhibited by
CC phosphoenolpyruvate.
CC -1- PATHWAY: Key control step of glycolysis.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the phosphofructokinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U51728; AAC45135.1; -.
CC EMBL: AL939111; CAB51967.1; -.
CC PIR: T35500; T35500.
CC HSSP: P00512; 3PFK.
CC HAMAP: MF_00339; -; 1.
CC InterPro: IPR000023; Pfructinase.
CC Pfam: PF00365; PFK, 1.
CC PRINTS: PRO0476; PHFCKTKINASE.
CC Prodom: PD000707; Pfructinase; 1.
CC PROSITE: PS00433; PHOSPHOFRUCTOKINASE; 1.
CC Allosteric enzyme; ATP-binding; Complete proteome; Glycolysis; Kinase;
KW Magnesium; Multigene family; Transferase.
FT NP_BIND 20 24 ATP (By similarity).
FT NP_BIND 155 159 ATP (By similarity).
FT NP_BIND 172 188 ATP (By similarity).
FT ACT_SITE 128 128 Proton acceptor (By similarity).
FT BINDING 163 163 Substrate (By similarity).
FT BINDING 266 266 Substrate (By similarity).
FT BINDING 272 272 Substrate (By similarity).
FT BINDING 275 275 Substrate (By similarity).
SQ SEQUENCE 342 AA; 36664 MW; CEEFC7B74092AB34 CRC64;

Query Match 13.6%; Score 311; DB 1; Length 342;
Best Local Similarity 27.4%; Pred. No. 1.3e-11;
Matches 105; Conservative 63; Mismatches 137; Indels 78; Gaps 17;

QY 34 KYAILTAGGLAPCLNSAIGSLIRYEIDPSLEITCYRGYKGLLGDSPVTAERYKKA 93
DB 2 KGVLTGGSDCCGLNAVITAVYRKGVQ-EYGVDFGFRGCMGPLEGDTVPL--DIPAR 58
QY 94 GYLORFGSGVIGNSRYKLTNVDCVKGGLVKEGEDPQKVAADQLVKDGVDILHTIGDDT 153
DB 59 GILPR-GGVLTASSSRNPLKORDGIR-----IKONLALGVEALITIGEDT 105
QY 154 NTRAADLAFLARNYGLTVIGLPTVNDVPEPIKQSLGAMTAAEGGARYFMNVVAENNA 213
DB 106 LGVATRLA-----DEGVGCVGPKTIDNLSATDTFDFDAVGATRAIDRLTTAAS 160
QY 214 NPEMLIVHEVMGNCMLTAATAGBYRKLLDRAEMLPGLGTRRESYEVAVFPEMAIDL 273
DB 161 HNRVLVY-EVMGRHACWILHS-----GLAGGA---NVILLPEGRFV 199
QY 274 EA-----EAKRLREVMKDVDCVNI FVSGAGVEAIVAEWQAGQEVPR---DAFGHIKLD 325

```

```

Db 200 EQVCSWTSRFRASVAPI-----VVVAEGA-----MPRODGMVLKDESLDSYGHVLS 247
QY 336 AVNPEKMFGEQPAQMITGAE-KTLVQSGYFAPASASNVDDMLIKSCADLAVECAFRRBS 384
Db 248 GV--GEMVLAKQIEKRTGNEARTTV--LGHVQSGTSPSAFDRWLARFGLHADV----- 298
QY 385 GVIGHDEENGV--LRATIEFPRI 405
Db 299 ---HGDGFGKVALRGTDIVRV 317

RESULT 9
PFP_DICTH STANDARD; PRT; 346 AA.
AC Q9K71;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pyrophosphate--fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.90)
DE (6-phosphofructokinase, pyrophosphate dependent) (pyrophosphate-
DE dependent 6-phosphofructose-1-kinase) (PPI-dependent
DE phosphofructokinase) (PPI-PFK).
GN Name=Pfp;
OS Dicyogloms thermophilum.
OC Bacteria; Dicyoglom; Dicyoglomales; Dicyoglomaceae; Dicyogloms.
OX NCBI_TaxID=14;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rt46 B.1;
RX MEDLINE=20372656; PubMed=10913106;
RX DOI=10.1128/JB.182.16.4661-4666.2000;
RA Ding Y.-H.R., Kontius R.S., Morgan H.W.;
RT "Sequencing, cloning, and high-level expression of the pfp gene,
RT encoding a ppi-dependent phosphofructokinase from the extremely
RT thermophilic eubacterium Dicyogloms thermophilum.";
RL J. Bacteriol. 182:4661-4666(2000).
CC -1- CATALYTIC ACTIVITY: Diphosphate + D-fructose 6-phosphate =
CC phosphate + D-fructose 1,6-bisphosphate.
CC -1- PATHWAY: Key control step of glycolysis.
CC -1- SIMILARITY: Belongs to the phosphofructokinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF268276; AAF80100.1; -.
DR HSSP; P06998; 2PFK.
DR HAMAP; MF_00339; -; 1.
DR InterPro; IPR000023; Ppfruckinase.
DR Pfam; PF00365; PFK; 1.
DR PROSITE; PS00476; PPFCKTKINASE.
DR PRINTS; PR000707; Ppfruckinase; 1.
DR ProDom; PD000707; Ppfruckinase; 1.
DR PROSITE; PS00433; PHOSPHOFRICTOKINASE; PALSE_NEG.
DR KX A10steric enzyme; ATP-binding; Glycolysis; Kinase; Magnesium;
DR transferase.
KM NP_BIND 23
FT NP_BIND 156 27 ATP (By similarity).
FT NP_BIND 173 160 ATP (By similarity).
FT NP_BIND 189 164 Proton acceptor (By similarity).
FT ACT_SITE 129 129 Substrate (By similarity).
FT BINDING 164 164 Substrate (By similarity).
FT BINDING 269 269 Substrate (By similarity).
FT BINDING 275 275 Substrate (By similarity).
FT BINDING 278 278 Substrate (By similarity).
SQ SEQUENCE 346 AA; 37448 MW; 50C03B64RA7927FI CRC64;

Query Match 13.4%; Score 305; DB 1; Length 346;
Beet Local Similarity 27.5%; Pred. No. 3e-11;
Matches 107; Conservative 73; Mismatches 155; Indels 54; Gaps 15;

```

```

QY 34 KVALITAGGLAPCLNSAIGSLIERYTEIPSIETIIICVAGGVKGLLGDSPYTAVERKKA 93
Db 5 RIGVLITGGDCPGINPARGIVMR--ADVDGEVIGLRYGNAGLLKADTMPISLEWED- 61
QY 94 GVLQFGGSSVIGNSRVXLTVNKDCVKGLYKEDGEPQKRAADQLVKQDVLITIGSDT 153
Db 62 --ILEIGTILGSSR-----TNP-----FKKEDEVQK-CVENPKKLLDLIALIGSDT 107
QY 154 NTAADLAFLARNNYGLTVIGLPRYDNDVFPPIKQSIGANTAAEQCARFENNVAAENNA 213
Db 108 LGVASKF-----SKLGIPMIGVPTIDKLEEDTYTGFPDAVEVVDAIKRLDTRAS 161
QY 214 NPMILVHEVMGRNCGMTAATAQEVRLDRAEMLPETGLTRSEYVHAIVFPMALD 273
Db 162 HARVIVV-ELINGRIAGLAL-----YGLAGADY-----ILIEVERNL 200
QY 274 EAEAKRLAEVMDK-VDCNVIFVSSAGVEALVAEMQAKGQVPRDPAFGIKLDAVNPGR 332
Db 201 EDLVNHRKLVARGRNNAVVAIAEGVQLPGFTYQ---KGQGMVDAFGHIRLGGV--GNV 255
QY 333 FGEQPAQMITGAEKTLVQSGYFAPASASNVDDMLIKSCADLAVECARRESGVIGHED 392
Db 256 LAEBIQKVLGIETRAVLIS-HLQKGGSPSIRDRIIMGLLGRKAVDLVHEKSGLF--VAV 312
QY 393 NGNVLRAIEFPRIKGGKPEIDTDMFNSM 421
Db 313 KGNELVPVDTILIG-KTRQNDVPAPFYESV 340

RESULT 10
PFP_AMEYE STANDARD; PRT; 341 AA.
AC Q59126;
DT 15-JUL-1998 (Rel. 35, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pyrophosphate--fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.90)
DE (6-phosphofructokinase, pyrophosphate dependent) (Pyrophosphate-
DE dependent 6-phosphofructose-1-kinase) (PPI-dependent
DE phosphofructokinase) (PPI-PFK).
GN Name=Pfp;
OS Amycolatopsis methanolica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardiaceae; Pseudonocardiaceae; Amycolatopsis.
OX NCBI_TaxID=1814;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96125240; PubMed=8550409;
RA Alves A.M., Meijer W.G., Vrijbloed J.W., Dijkhuizen L.;
RT "Characterization and phylogeny of the pfp gene of Amycolatopsis
RT methanolica encoding ppi-dependent phosphofructokinase.";
RL J. Bacteriol. 178:149-155(1996).
CC -1- CATALYTIC ACTIVITY: Diphosphate + D-fructose 6-phosphate =
CC phosphate + D-fructose 1,6-bisphosphate.
CC -1- PATHWAY: Key control step of glycolysis.
CC -1- SUBUNIT: Homotrimer.
CC -1- SIMILARITY: Belongs to the phosphofructokinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U31277; AAB01683.1; ALT_INIT.
DR HSSP; P00512; 3PFK.
DR HAMAP; MF_00339; -; 1.
DR InterPro; IPR000023; Ppfruckinase.
DR Pfam; PF00365; PFK; 1.
DR PRINTS; PR00476; PPFCKTKINASE.
DR ProDom; PD000707; Ppfruckinase; 1.

```


DR PROSITE, PS00433; PHOSPHOFRUCTOKINASE, 1.
 KW Allosteric enzyme; ATP-binding; Direct protein sequencing; glycolysis;
 KM Kinase; Magnesium; Transferase.
 FT NP_BIND 20 24
 FT NP_BIND 154 158
 FT NP_BIND 171 187
 FT ACT_SITE 127 127
 FT BINDING 162 162
 FT BINDING 265 265
 FT BINDING 271 271
 FT BINDING 274 274
 SQ SEQUENCE 341 AA; 36229 MW; 30661346172D5B CRC64;
 Query Match 13.1%; Score 298.5; DB 1; Length 341;
 Best Local Similarity 28.8%; Pred. No. 7.7e-11;
 Matches 108; Conservative 59; Mismatches 147; Indels 61; Gaps 16;
 QY 34 KVAITLGGALCLNSAIGSLIERYTEIDPSIEIICYRGYKGLIGDSYPYTAAYRKA 93
 DB 2 RVGVLTGGDCGCLNANAVIRGVKGIENH-GWEIVGFRSGMGPLTGDSPRLGLDVEE- 59
 QY 94 GVLQRFSGSVGNRSVKLTNVKDCVRGLVKGEDPQKVAADQLVKDGLDHTTGGDDT 153
 DB 60 -ILIR-GGITLSSSR---TNP-----YKEBGVEKIRA-VLADQGVDAIATIGEDT 105
 QY 154 NTAADLAFLARNYGLTVIGLPTVDNDVPEIKOSLGAMTPAEGARFMMVVAENNA 213
 DB 106 LGVAKKL-----TDGIGVGVPTKIDVDLADTDFPDPAVHATAIIRLTAA 159
 QY 214 NBRMLIVHEVGNRCGMLTAATAQERYKLLDRAEWLPELIGLTRESYEVHAFVPEMAIDL 273
 DB 160 HYRAWVW-EVMGRHAGMIAL-----HAGLAGGA---NVILPEPFSV 198
 QY 274 EBAEKRLREVMKVCNITFVSEGAIVEAEMOKAGEPRDPAAGHILKDAVNGKRF 333
 DB 199 EGVVEVERRRFKMAPRIIVABGAVPEG-GAEVLRTGE---KDAIGHQVLGCV-GTWL 252
 QY 334 GEOPAFMIGAEKTLVQKSGYFARASANDMRLIKSCADLAVECAFRESGVITGHEDN 393
 DB 255 ADEIABRTCKESRAV-VLGHTRGCTPTAYDRVLARFGLHVDAN-----ADGDF 302
 QY 394 GNV--LRAIEFPRIK 406
 DB 303 GTMVALLRGTDIVRVK 317
 RESULT 11
 Q82AH4 PRELIMINARY; PRT; 342 AA.
 AC 082AH4;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Putative 6-phosphofructokinase.
 GN Name=pfkA3; OrderedLocustNames=SAV6083;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxId=33903;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=WA-4680;
 RX MEDLINE=2608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 Sakaki Y., Hattori M., Omura S.,
 RT "Complete genome sequence and comparative analysis of the industrial
 microorganism Streptomyces avermitilis.";
 RL Nat. Biotechnol. 21:526-531(2003).
 RN (12)
 RP SEQUENCE FROM N.A.
 RC STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
 RT "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: deducing the ability of producing secondary
 metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 CC -1- PATHWAY: Key control step of glycolysis.
 DR EMBL, AP005045; BAC73794.1; --
 DR HSSP, P00512; 3PFK.
 DR GO, GO:0005945; C-6-phosphofructokinase complex; IEA.
 DR GO, GO:0003872; P-6-phosphofructokinase activity; IEA.
 DR GO, GO:0006096; P-glycolysis; IEA.
 DR InterPro, IPR000023; Pfructkinase.
 DR Pfam, PF00365; PFK, 1.
 DR PRINTS, PR00476; Pfructkinase.
 DR PRODOM, PD000707; Pfructkinase.
 DR PROSITE, PS00433; PHOSPHOFRUCTOKINASE, 1.
 KW Complete proteome; Glycolysis; Kinase; Transferase.
 SQ SEQUENCE 342 AA; 36724 MW; 5A2C3FA82917CCDF CRC64;
 Query Match 12.9%; Score 294; DB 2; Length 342;
 Best Local Similarity 26.5%; Pred. No. 1.5e-10;
 Matches 101; Conservative 66; Mismatches 140; Indels 74; Gaps 17;
 QY 34 KVAITLGGALCLNSAIGSLIERYTEIDPSIEIICYRGYKGLIGDSYPYTAAYRKA 93
 DB 2 RVGVLTGGDCGCLNANAVIRGVKGIQV-BYGVDPGFRGMGPLEGDA--VRLDPAVR 56
 QY 94 GVLQRFSGSVGNRSVKLTNVKDCVRGLVKGEDPQKVAADQLVKDGLDHTTGGDDT 153
 DB 59 GILPR-GGITLSSSRTPKLDDGIR--IKE-----NLAKQEVDAIATIGEDT 105
 QY 154 NTAADLAFLARNYGLTVIGLPTVDNDVPEIKOSLGAMTPAEGARF--MMVVAEN 211
 DB 106 LGVAKRLT-----DEGVAVVGVPTKIDVDLSATDTFEPDPAVGIATAIIRLTAA 160
 QY 212 NANPMLIVHEVGNRCGMLTAATAQERYKLLDRAEWLPELIGLTRESYEVHAFVPEMAI 271
 DB 161 HMR---VLCEVMGNHAGIAIHS-----GLAGGA---NVILPEQGF 197
 QY 272 DIEBAEKRLREVMKVCNITFVSEGAIVEAEMOKAGEVPR---DAFGHILKDAV 327
 DB 198 DVDQVCAYVTSRFRKASYAPIVVAEGA-----MPKQGVNVLKDESLDSFGHVALSGV 249
 QY 328 NPGKWFGEOPAFMIGAE-KTLVQKSGYFARASANDMRLIKSCADLAVECAFRESGV 386
 DB 250 --GEWLAKIEIKRTGEARTTV--LGHVORGCTPSAFDRMLATRFGLHAIEAV----- 298
 QY 387 IGHDEDNGNV--LRAIEFPRI 405
 DB 299 --RQDGRGMVALRGTDIVRV 317
 RESULT 12
 Q82603 PRELIMINARY; PRT; 341 AA.
 AC 082603;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Putative 6-phosphofructokinase.
 GN Name=pfkA4; OrderedLocustNames=SAV7123;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxId=33903;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=WA-4680;
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,

RX MEDLINE:21996410; PubMed:12000953; DOI=10.1038/417141a;
RA Bentley S.D., James K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., Chater K.D., Harris D.E., Quail M.A., Kleiser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleiser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)";
RL Nature 417:141-147(2002).
CC -1- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
CC fructose 1,6-bisphosphate.
CC -1- PATHWAY: Key control step of glycolysis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the phosphofructokinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF39108; CAC01496.1; -
DR HSSP: P00512; 3PFK. -
DR HAMAP: MF_00339; 1.
DR InterPro: IPR000023; Pfkfructkinase.
DR Pfam: PF00365; PFK, 1.
DR PRINTS: PR00476; Pfkfructkinase.
DR PRODOM: PD000707; Pfkfructkinase; 1.
DR PROSITE: PS00433; PHOSPHOFRUCTOKINASE; 1.
DR Allotopic enzyme; ATP-binding; Complete proteome; Glycolysis; Kinase;
KW Magnesium; Multigene family; Transferrase.
FT NP_BIND 20 24 ATP (By similarity).
FT NP_BIND 154 158 ATP (By similarity).
FT NP_BIND 171 187 ATP (By similarity).
FT ACT_SITE 127 127 Proton acceptor (By similarity).
FT BINDING 162 162 Substrate (By similarity).
FT BINDING 266 266 Substrate (By similarity).
FT BINDING 272 272 Substrate (By similarity).
FT BINDING 275 275 Substrate (By similarity).
SQ SEQUENCE 341 AA; 36431 MW; E3050D37BDB6P9F0 CRC64;

Query Match 12.3%; Score 281.5; DB 1; Length 341;
Best Local Similarity 26.8%; Pred. No. 9e-10;
Matches 100; Conservative 66; Mismatches 152; Indels 55; Gaps 12;

QY 34 KAILTLAGLAPLCLNSAIGSLERYTEIDPSEIILCYRGYGLLGDSPYTAERKKA 93
DB 2 RIGVLISGGDCPELNAVIRSVVHR-AVHDGDEVIFRGMKGLCDY-ELDLDAVG 58
QY 94 GVLQREGSVIGNSRYKLTNVKDCVKGEGEDPQKVAADQLVKGVDYLLHTIGDDT 153
DB 59 GILAR-CGIIILSSRRRPHRLRGVER-----ARGHVEELGIDAIIPIGSEGT 105
QY 154 NTAADLAFLARNNYGLTVIGLPTVDNDVFPKISLGAMTAEGARFPMNVVAENNA 213
DB 106 LKAARLL-----SDNGLPIVGPKTIIDNDIATDVTGFDPAVTATBALDLKTTAS 159
QY 214 NPMRLIVHEVMGRNCMLTAATQAEYRKLLDRAEWLPELGLTRESVEYNAVPEMAIDL 273
DB 160 HORVLV-LVEMGHMTWIALHSGM-----AAGAAVAVVPERPFI 198
QY 274 EAEAKLREVMQKVCVNIFFV-SEGAGVEAIVAEQAKQGEVPRDFGHIKLDVAVPGKM 332
DB 199 DELTAVGRFSAKGRFATVVAEGAIPKAGTMDPEGG-----KDYVGHERRFGI--AQ 252
QY 333 FEGQFQMTIGAELTVQKGYFAPASASVNDMLIKSCADLAVECAFRESGVI---G 388

DB 253 LSEIEERLKEARPV-ILGHVQGGTPTAYRVLATRFMGHVAEVAHGEFGKMTALNG 311
QY 389 HDENGNVRARIE 401
DB 312 TDIEMVSLADAVE 324

RESULT 15
ID 09CSJ7 PRELIMINARY; PRT; 485 AA.
AC 09CSJ7;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Putative pyrophosphate-dependent phosphofructo-1-kinase.
GN Name=At5g56630;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Setou M.,
RA Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Egu P., Lee J.M.,
RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shim P.,
RA Southwick A., Davis R.W., Ecker J.R., Theologis A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF360207; AK25917.1; -
DR EMBL: AY040055; AK64113.1; -
DR HSSP: P00512; 1MTO.
DR GO: GO:0005945; C:6-phosphofructokinase complex; IEA.
DR GO: GO:0008872; F:6-phosphofructokinase activity; IEA.
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:006096; P:glycolysis; IEA.
DR InterPro: IPR000023; Pfkfructkinase.
DR Pfam: PF00365; PFK, 1.
DR PRINTS: PR00476; Pfkfructkinase.
DR PRODOM: PD000707; Pfkfructkinase; 1.
KW Kinase.
SQ SEQUENCE 485 AA; 53482 MW; 6F0C6CF93EAA7B5 CRC64;

Query Match 12.3%; Score 281.5; DB 2; Length 485;
Best Local Similarity 26.7%; Pred. No. 1.4e-09;
Matches 109; Conservative 61; Mismatches 166; Indels 73; Gaps 16;

QY 37 IITAGLAPLCLNSAIGSLERYTEIDPSEIILCYRGYGLLGDSPYTAERKKA 96
DB 96 IYCGSLCPGLNTVIREVSSLSYMGVKKIIGDGGHGFPAKTIPLNSKV--VNDI 152
QY 97 QREGSVIGNSRYKLTNVKDCVKGEGEDPQKVAADQLVKGVDYLLHTIGDDTNTA 156
DB 153 HKRGITIGTSR-----CGHDTNNTI-VDSIDRGINQVYIIGGDTQNG 195
QY 157 AADLAFLARNNYGLTVIGLPTVDNDVFPKISLGAMTAEGARF--MNVVAENNA 214
DB 196 ASVIFEEIRRRRLKAAVVGIPPTINDIPVDSKSGFDPAVEAQRATNAHVEAESNEN 255
QY 215 PMRLIVHEVMGRNCMLTAATQAEYRKLLDRAEWLPELGLTRESVEYNAVPEMAIDL 274
DB 256 GIGFV--KUMGRYSGIIA-----MATLASRDVDCCLIPSEFFYLE 294
QY 275 AEA-----KRLREVMQKVCVNIFFVSEGAGVEAIVAEQAKQGEVPRDFAGHIKLDV 327

```

Db      295 GEGGLPEPTERRLKDHGHNV---IVLAEGAGQDLMCKSMES---TPMDASGNKLLKDV 346
Oy      328 NPGKNFGQFAQMIGAEKTLVQ---KSGYFARASASNVDDRLIKSCADLAVECAFRRE 383
Db      347 --GLMSQSIKDHFKKKNVMMLKXIDPTYMIRAVPSNASDNYCTLLAQSAVHGAM--- 401
Oy      384 SGVIGHDED--NGNVLRAIEPPRIKGGKPFENIDTD--WNSMLSEIGQP 428
Db      402 AGYGTSGLVNGR-QTYIPFYRITETQNNVVITDRMMAR-LLSSTNCP 448

```

Search completed: February 24, 2005, 14:22:28
 Job time : 185 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 14:01:54 ; Search time 40 Seconds
(without alignments)
1051.168 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTWPHYLTDIRFCWHFF.....FNSMSEIQPKGKRVESH 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1314	57.5	405	2 A41169	diphosphate-fructo
2	1292.5	56.6	479	2 AC2836	hydrophobic prote
3	1292.5	56.6	514	2 P97613	ppl-phosphofructok
4	311	13.6	342	2 T35500	6-phosphofructokin
5	270	11.8	448	1 F70190	probable diphospha
6	263.5	11.5	436	1 S49458	diphosphate-fructo
7	257.5	11.3	500	2 T06011	probable diphospha
8	253	11.1	319	2 C72406	6-phosphofructokin
9	251	11.0	473	2 T13433	pyrophosphate-depe
10	246.5	10.8	987	1 JQ0016	6-phosphofructokin
11	245.5	10.7	992	2 S32902	6-phosphofructokin
12	243	10.6	462	2 T10691	probable diphospha
13	241.5	10.6	437	2 S54978	6-phosphofructokin
14	240.5	10.5	320	2 AH0942	6-phosphofructokin
15	238.5	10.4	320	1 KIECPA	6-phosphofructokin
16	238.5	10.4	320	2 AB6081	6-phosphofructokin
17	238.5	10.4	320	2 A98234	6-phosphofructokin
18	234.5	10.3	327	2 AF0010	6-phosphofructokin
19	233.5	10.2	322	2 S39248	6-phosphofructokin
20	233	10.2	319	2 S35928	6-phosphofructokin
21	232	10.2	461	1 A71366	diphosphate-fructo
22	226	9.9	322	1 JQ1028	6-phosphofructokin
23	224.5	9.8	334	2 A75495	6-phosphofructokin
24	223.5	9.8	319	2 D84045	6-phosphofructokin
25	223.5	9.8	331	2 G82045	6-phosphofructokin
26	223	9.8	343	2 E70856	probable pfka prot
27	222	9.7	340	1 JN0614	6-phosphofructokin
28	221	9.7	335	2 D97971	6-phosphofructokin
29	220	9.6	573	2 C71312	probable pyrophosp

30	220	9.6	780	1 KIRBF	6-phosphofructokin
31	216.5	9.5	319	2 P96963	6-phosphofructokin
32	215	9.4	335	2 P95103	6-phosphofructokin
33	213.5	9.3	321	2 C70447	phosphofructokinase
34	211.5	9.3	348	2 T45407	phosphofructokinase
35	210.5	9.2	320	2 P84965	6-phosphofructokin
36	210.5	9.2	343	2 G87121	6-phosphofructokin
37	210.5	9.2	544	2 S52081	diphosphate-fructo
38	210	9.2	780	1 KIRHFM	6-phosphofructokin
39	210	9.2	789	1 A45617	6-phosphofructokin
40	209	9.2	319	1 KIRBFF	6-phosphofructokin
41	209	9.2	323	1 G64223	6-phosphofructokin
42	207.5	9.1	791	2 A53206	6-phosphofructokin
43	207	9.1	959	1 JQ2015	6-phosphofructokin
44	204.5	9.0	784	2 JC2057	6-phosphofructokin
45	204	8.9	463	2 B84613	hypothetical prote

ALIGNMENTS

RESULT 1

A41169
diphosphate-fructose-6-phosphate-1-phosphotransferase (EC 2.7.1.90) - Propionibacterium f
N:Alternate names: 6-phosphofructokinase (pyrophosphate)
C:Species: Propionibacterium freudenreichii
C>Date: 17-Jul-1992 #sequence_revision 15-Mar-2004 #ext_change 09-Jul-2004
C/Accession: A41169
R:Radnor, U.S.; Gollapudi, L.; Tripathi, R.L.; Latschaw, S.P.; Kemp, R.G.
U. Biol. Chem. 266, 16550-16555, 1991
A/Title: Cloning, sequencing, and expression of pyrophosphate-dependent phosphofructokin
A/Reference number: A41169; PMID:9158443; PMID:1653240
A/Molecule type: DNA
A/Residues: 1-404 <LAD>
A/Cross-references: UNIPROT:P29495; GB:M67447; NID:g150930; PIND:AAA2675.1; PID:g150931
C/Comment: This is an example of a nonallosteric, pyrophosphate-dependent phosphofructoki
C/Superfamily: pyrophosphate-dependent phosphofructokinase, SMC01852 type
C/Keywords: phosphotransferase
F:5-322/Domain: 6-phosphofructokinase 1 homology <6PFI>

Query Match	Score	1314;	DB 2;	Length	405;
Best Local Similarity	62.6%;	Pred. No. 7.8e-85;			
Matches	248;	Conservative	60;	Mismatches	88;
				Indels	0;
				Gaps	0;
QY	33	KKVAITTAGGLAPCNSAIGSLIERTEIDPSLEICVGGYKGLIGSPYTAVERKK	92		
DB	4	KKVALTTAGGAPACUSSAIAELIKRYTEVSPETTLIGRYGEGLLKGSLEFSPAVRAH	63		
QY	93	AGVLQRFQGSVIGNSRVKLTNVKDCVKGVLVEGEDPQKVAADQVLKDGVDILHTIGDD	152		
DB	64	YRLRFSGSPIGNSVKLTNVKDLVARGLVASGDPLKVAADQVLADGVDVHTIGDD	123		
QY	153	TNTAAADLAFLARNNGYLTGTPPTVDNDVFPKISIGAMTAAGARFNNVAENN	212		
DB	124	TNTTAADLAAYLAQHDYPLTVVGLPKTINDIYPIRISIGAMTAADGARFANNVAENN	183		
QY	213	ANPRMLIVHEWNGRNGLTAATAQERYKLDRAEKLPELGLTREYVHAHFVPEMAD	272		
DB	184	AAAPRELIIHEIRNGRCGYLAATSRRYVAMLAQOQLPBAQDRGMDIHALYVPEATID	243		
QY	273	LEAEAKRLREWMDKVDGVNI FVSEGAGVEAIVAEKMAQGEVPRDAFGIKLDAVPGKW	332		
DB	244	LDAEAKRLTVDVDSVNI FVSEGAGVDIYAQMOTQEVPTDAFGVQDKINPGAW	303		
QY	333	FEQQAQMIAGKTLVQKSGYFAPASASVNDWMRLIKSCADLAIVECAFRRREGVIGHDD	392		
DB	304	FAKQFAERIGAKTWQKSGYFSRSKSAQOLELTAATVAATVAADLAAGTGGVGDDBE	363		
QY	393	NGNVLAIEPRPIKGGKPNIDTDWNSMLSEIQG	428		
DB	364	AGDKLSVIDFKRIAGHKPPDITLDWYTQLARIQGP	399		

RESULT 2
AC2836
hypothetical protein pfp [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #ext_change 09-Jul-2004
C/Accession: AC2836
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Genthner, D.; Kuyavlin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Author: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
er, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AC2836
A/Status: preliminary
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-479 <KUR>
A/Cross-references: UNIPROT:Q8UDL0; GB:AB008688; PIDN:AAL43105.1; PID:g17740576; GSPDB:G
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Map position: circular chromosome
C/Superfamily: pyrophosphate-dependent phosphofructokinase, SMC01852 type; 6-phosphofruc
Query Match 56.6%; Score 1292.5; DB 2; Length 479;
Best Local Similarity 57.9%; Pred. No. 3.2e-83;
Matches 246; Conservative 71; Mismatches 101; Indels 7; Gaps 2;
OY 8 HLTADI-----RCHMFILNFNYTLNKKKVAITLTAAGLAPCLNSAIGSLIERTIETDP 63
DB 54 HLHAAVASICQTCCLWVYAK---RRAMMAKQVAMLTAGLAPCLSSAVGLIERYSDIAP 110
OY 64 SIEITCYRGYKGLLDGDSYPTAEVRKKAAGVLOFRGSSVIGNSRVKLTNNVDCYKRGIV 123
DB 111 EIDITAYRSGYGVLLGRIETITKMRKALHLHRYGSSPIGNSRVKLTNADCAKRGIV 170
OY 124 KEGEDPOKVAADQVLDGVDILHTIGDDTNTAAADLAFLARNNYGLTVIGLPTVDND 183
DB 171 KEGDNPRLVAERLAADGITLHTIGDDTNTAAADLAFLANGYDLTVGLPTVDND 230
OY 184 VEPITQSLGATTAEOGARFPMNVVAENNANPRMLIVHEVMGRNGMLTAATAOERYKL 243
DB 231 VVPIKQSLGATTAEGVGAFFDNVSNESQAPKTFVIEHVMGRHCGMLTAATARAAYIOKT 290
OY 244 DRAEWLPGLTRESYEVAHVPEMAITDLAEAKRLREVMDCVNVFVSEGAIVEAI 303
DB 291 RENDYVEGLMMNTQKNIGITLPEMAFDIEAEERLKEVMDKHGYVTLFVSEGAIDLAI 350
OY 304 VAEMQAKQGEVPRDAFGHIKLDVAVPGKMFGEQFQAGTIVYOKSGYFARASAVND 363
DB 351 VAREBAAGVAKRDAFGVHKIDITINVGWFOQFQAGLGAERSVQKSGYFARASAPAND 410
OY 364 DMRLIKSCADLAVECAFRRBSGVIGHDENGVLAIEFPRIKGGKFPNIDTWMSMLS 423
DB 411 DLRLIQWVDLAVESALKVSGVGHDEDQNGKLTIEFPRIKGGKHFDLSAKWFTVEWE 470
OY 424 BIGQP 428
DB 471 HVGP 475

RESULT 3
P97613
pol-phosphofructokinase (AF246209) [imported] - Agrobacterium tumefaciens (strain C58, C
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #ext_change 09-Jul-2004
C/Accession: P97613
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2333-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: P97613
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-514 <KUR>
A/Cross-references: UNIPROT:Q8UDL0; GB:AB007869; PIDN:AAK87863.1; PID:g15157249; GSPDB:G
C/Genetics:
A/Map position: circular chromosome
C/Superfamily: pyrophosphate-dependent phosphofructokinase, SMC01852 type
Query Match 56.6%; Score 1292.5; DB 2; Length 514;
Best Local Similarity 57.9%; Pred. No. 3.5e-83;
Matches 246; Conservative 71; Mismatches 101; Indels 7; Gaps 2;
OY 8 HLTADI-----RCHMFILNFNYTLNKKKVAITLTAAGLAPCLNSAIGSLIERTIETDP 63
DB 89 HLHAAVASICQTCCLWVYAK---RRAMMAKQVAMLTAGLAPCLSSAVGLIERYSDIAP 145
OY 64 SIEITCYRGYKGLLDGDSYPTAEVRKKAAGVLOFRGSSVIGNSRVKLTNNVDCYKRGIV 123
DB 146 EIDITAYRSGYGVLLGRIETITKMRKALHLHRYGSSPIGNSRVKLTNADCAKRGIV 205
OY 124 KEGEDPOKVAADQVLDGVDILHTIGDDTNTAAADLAFLARNNYGLTVIGLPTVDND 183
DB 206 KEGDNPRLVAERLAADGITLHTIGDDTNTAAADLAFLANGYDLTVGLPTVDND 265
OY 184 VEPITQSLGATTAEOGARFPMNVVAENNANPRMLIVHEVMGRNGMLTAATAOERYKL 243
DB 266 VVPIKQSLGATTAEGVGAFFDNVSNESQAPKTFVIEHVMGRHCGMLTAATARAAYIOKT 325
OY 244 DRAEWLPGLTRESYEVAHVPEMAITDLAEAKRLREVMDCVNVFVSEGAIVEAI 303
DB 326 RENDYVEGLMMNTQKNIGITLPEMAFDIEAEERLKEVMDKHGYVTLFVSEGAIDLAI 385
OY 304 VAEMQAKQGEVPRDAFGHIKLDVAVPGKMFGEQFQAGTIVYOKSGYFARASAVND 363
DB 366 VAREBAAGVAKRDAFGVHKIDITINVGWFOQFQAGLGAERSVQKSGYFARASAPAND 445
OY 364 DMRLIKSCADLAVECAFRRBSGVIGHDENGVLAIEFPRIKGGKFPNIDTWMSMLS 423
DB 446 DLRLIQWVDLAVESALKVSGVGHDEDQNGKLTIEFPRIKGGKHFDLSAKWFTVEWE 505
OY 424 BIGQP 428
DB 506 HVGP 510

RESULT 4
T35500
6-phosphofructokinase (EC 2.7.1.11) - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #ext_change 09-Jul-2004
C/Accession: T35500; T42063
R/Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A/Reference number: 221580
A/Accession: T35500
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-342 <SES>
A/Cross-references: UNIPROT:O08333; EMBL:AL109661; PIDN:CAB51967.1; GSPDB:GN00070; SCOE
A/Experimental source: strain A2(2)
R/Alves, A.M.; Euvierink, G.J.; Bibb, M.J.; Dijkhuizen, L.
Appl. Environ. Microbiol. 63, 956-961, 1997
A/Title: Identification of ATP-dependent phosphofructokinase as a regulatory step in the
A/Reference number: 222050; MUID:97208211; PMID:9055413
A/Accession: T42063
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-342 <ALV>
A/Cross-references: EMBL:U51728; NID:g1931572; PIDN:AA45135.1; PID:g1931573
C/Genetics:


```

QY 212 NANPMLIVHEVMGRNCGLTAATAOYRKLIDRAEWLPELGTRESYEVAVPEMAI 271
DB 242 KSAKKGIGIVRLMGHDAGFIAL-----YASLANG-----DANLVLIPEIDI 282
QY 272 DL-----EAEAKRLREVMDVDCNIFVSEAGVEAIVAEOMAKGQEPDPAFGHI-KIDA 326
DB 283 PFTQICEFPGKR-----IMSKGHV-IVVAEGA-----LONQKRDLDLGTSGNLIHDS 333
QY 327 VNPGRFGEQFAQMIAGKTLVQ--KSGYFAPASASNVDMRLIKSCADLAVECAFRRRS 384
DB 334 INVLDSITKYLKISGIEHTIKFVDPSPYMRSPACSAADAFHMCMLANAAVHVAMAGKT 393
QY 385 G-VIGHDENGVNLAIEFPR-----IKGKRPNIIDDM-FNSMLSEIQPK 429
DB 394 GLVICHNNHNN---FVSPIDRTSYIK-----RVMTDGLYTMASAIERPK 436

RESULT 7
T06011
Probable diphosphate-fructose-6-phosphate 1-phosphotransferase (EC 2.7.1.90) T25K17.80
N/Alternate names: protein T25K17.80
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
R/Beyan, M.; Rose, M.; Hempel, S.; Entlan, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, March 1999
A/Accession: T06011
A/Reference number: Z15382
A/Accession: T06011
A/Molecule type: DNA
A/Residues: 1-500 <BEV>
A/Cross-references: UNIPROT:Q9SMQ7; EMBL:AL049171; GSPDB:GN00062; ATSP:T25K17.80
A/Experimental source: cultivar Columbia; BAC clone T25K17
A/Map position: 4
A/Genes: ATSP:T25K17.80
A/Intons: 46/2; 59/3; 80/3; 141/3; 194/3; 212/3; 235/3; 279/1; 367/3; 383/1; 434/3
C/Superfamily: pyrophosphate-dependent phosphofructokinase, TP0108 type; 6-phosphofructo
C/Keywords: phosphotransferase

Query Match 11.1%; Score 257.5; DB 2; Length 500;
Best Local Similarity 25.6%; Pred. No. 1.9e-10;
Matches 109; Conservative 59; Mismatches 167; Indels 91; Gaps 16;

QY 37 ILTAGLAPCLNSAIGSL-----IERYTEIDPSIEIIC-----YRGYKGL 78
DB 92 IVTCGLCPGLMTVIREIVSSLSVMGYKRIIGIDVS---FCDCNLLLTKTNTGGRGY 148
QY 79 LGDSYPTVAEVRKKAQVLRFGSGVIGNSRVLTIVKDCVKGGLVKEGEDPQKVAADQV 138
DB 149 AKNTVSLDSKV---VNDIHKRGITLGTSR-----GGHDTTKI-VDSIQ 188
QY 139 KDGVDIHLITGGDNTTAADLAFLARNNYGLTVIGPKTVNDVFPKISGLGAMTAAE 198
DB 189 DRGIVQVYITGGDGRGASVIFEEIRRRGLKVAIVIGIPKTIIDNIPIVIDSFGDTAVE 248
QY 199 QGARYFMNVAVENNANPMLIVHEVMGRNCGLTAATAOYRKLIDRAEWLPELGTRES 258
DB 249 EAQRIANNAHVAEBSIENGIGVVKLMGRYSFIA-----MYATLAS 289
QY 259 YEVAHVPEMAIDLEAEA-----KRLREVMKVDCNIFVSEAGVEAIVAEOMAKG 311
DB 290 RDVDCCLIPESPFIYEGEGGLEFYIEKRLKESGHNV---LVIAAGAGDGLMSKMSHES- 344
QY 312 QEVPDAGHILKLDVNPGRFGEQFAQMIAGKTLVQ--KSGYFAPASASNVDMRL 367
DB 345 --TLMDASGNKLLKQV--GLMISQSIKHFNOKKVMNLKTYDIPYMRRAVSNASNDVY 400
QY 368 IKSADLAVECAFRRBSGVTGHDENGVN--LRAIEPRIKGKRPNIIDDM-FNSML 422
DB 401 CTLLAQSAVHGAMAGVTGYI-----SGLVNGRQTYIPFYRIKONHVVITDRMMAR-L 454
QY 423 SEIQPK 428

```

```

DB 455 SSTNQP 460

RESULT 8
C72406
6-phosphofructokinase - Thermotoga maritima (strain MS8)
C/Species: Thermotoga maritima
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: C72406
R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nucleotide 399, 323-329, 1999
A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A/Reference number: A72200; MUID:99287316; PMID:10360571
A/Accession: C72406
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-319 <ARN>
A/Cross-references: UNIPROT:Q9WY52; GB:AE001705; GB:AE000512; NID:94980694; PIDN:AA035301
A/Experimental source: strain MS8
C/Genetics:
A/Genes: TM0209
C/Superfamily: ATP-dependent phosphofructokinase, prokaryotic type; 6-phosphofructokinase
F/4-279/Domain: 6-phosphofructokinase 1 homology <6FP>

Query Match 11.1%; Score 253; DB 2; Length 319;
Best Local Similarity 26.6%; Pred. No. 2.1e-10;
Matches 95; Conservative 60; Mismatches 126; Indels 76; Gaps 14;

QY 33 KKVAILTNGSLAPCLNSAIGSLIERYTIDISIEICTRGYKGLLDSDSYPTVAEVRK 92
DB 2 KKIIVLTSGDAPGNNAVAVV-RY-GVRQGLEVIQVRGSGIIDDF--VKLEYKDV 57
QY 93 AGVLRFGSGVIGNSRVLTIVKDCVKGGLVKEGEDPQKVAADQVKGVDIHLITGGD 152
DB 58 AGITRK-GGITLRISRCFEFT-----EEGELAAQIKKKGIEGLVIGSEG 104
QY 153 TTTAAADLAFLARNNYGLTVIGPKTVNDVFPKISGLGAMTAAEQARFPMNVAVENN 212
DB 105 SLT-----GAILLVEBHKIPVVGIPATIDNDIGLDMCIGVDTCLNTWDAVQKLDRTAS 159
QY 213 ANPMLIVHEVMGRNCGLTAATAOYRKLIDRAEWLPELGTRESYEVAVPEMAID 272
DB 160 SHERAFIV-EVMGRHSGVIALMAG---LVTGAE-----AIYVEIPVD 198
QY 273 LEAEAKRLREV--MDKVCNVI FVSEAGVEAIVAEOMAKGQEPDPAFGHIKLDVAVNP 330
DB 199 YSQLADRILIEERRKINSI-IVAEAGASAYTVAR-----HLE----- 236
QY 331 KMFGEQFAQMIAGKTLVQKSGYFAPASASNVDMRLIKSCADLAVECAFRRSGVI 387
DB 227 -----YRIGYE-TRTITIGHVQGRGSPAPAFDRRLLSHGVEAVDALLDQEVVDM 284

RESULT 9
T13433
pyrophosphate-dependent phosphofructo-1-kinase homolog T17A13.40 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C/Accession: T13433
R/Beyan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft
submitted to the Protein Sequence Database, July 1999
A/Reference number: Z17683
A/Accession: T13433
A/Molecule type: DNA
A/Residues: 1-473 <BEV>
A/Cross-references: UNIPROT:Q9WOF9; EMBL:AL096692; GSPDB:GN00062; ATSP:T17A13.40
A/Experimental source: cultivar Columbia; BAC clone F17A13
A/Genes: ATSP:T17A13.40
A/Map position: 4

```


A.Introns: 35/3; 51/2; 64/3; 85/3; 131/3; 202/3; 225/3; 269/1; 358/3; 373/1; 424/3
 C/Superfamily: pyrophosphate-dependent phosphofructokinase, TP0108 type; 6-phosphofructo
 F,95-405/Domain: 6-phosphofructokinase 1 homology <6PF>

Query Match 11.0%; Score 251; DB 2; Length 473;

Best Local Similarity 25.2%; Pred. No. 5e-10; Mismatches 174; Indels 72; Gaps 15;

Matches 103; Conservative 60; Mismatches 174; Indels 72; Gaps 15;

QY 37 ILTLAGLAPCLNSALISLERYEIDPSIEIICRGYKGLLDGSPYTAERKKAGYL 96
 DB 97 ITCGGLCGMLNTVREIYCGLSYMGVGRILIGDGYGFAARNTHIDL---KTVNDI 153
 QY 97 QRFSGSVGNRSVKLTNVKDCVKRGLVKEGDEPOKVAADQLVKDGVDILHTTGDDPTNTA 156
 DB 154 HRSGGTILGTSTR-----GGHNTTKI-VSIOQRGINQVYIIGGDSQKXG 196
 QY 157 AADLAFLARNYGLTVIGPKTVNDVFPKISGAMTAEGAGYFNNVNAENNANR 216
 DB 197 AALIFEIRKRLKLVAVAGIPKTIIDIPIDRSFGPDVAEQAQAINAAHVEATSPEN 256
 QY 217 MLIVEMGRNCGMLTAATAQERYKLLDBAEMLPBGLTRSEYVAHVPEMAIDLEAE 276
 DB 257 GGLVLMKRYGFI-----WATLASRDVCCILPESPFLBGS 297
 QY 277 A-----KRLBWDKVDVCNIFVSEGAVEAIVEMQAKGEVPRDAFGHILDAVNP 329
 DB 298 GGLFEFIDRLKESGHV-----IVIEAGAGDLSSESM---KESTTLKDSGNLTLDI-- 349
 QY 330 GKPFQGFQMGAEKTLVOK---SGYFAPASNSNDVRLKSCADLAVECFRRESGV 386
 DB 350 GLMISRIKDHAKKWTLLTKTIDPTMYRAVPSNSD---NVCTLLAQSAVH---GV 402
 QY 387 I-GHDEDNNGVLA---IEPPI-KGKGFNIIDWFMSLSEIGOP 428
 DB 403 MAGYNGFTYGLVNGRHTYIPFNRIKOKKVVITDMMAR-LLSTNQF 450

RESULT 10

J00016 6-phosphofructokinase (EC 2.7.1.11) alpha chain - yeast (*Saccharomyces cerevisiae*)

N.Alternate names: phosphofructokinase 1; phosphohexokinase; protein 194; protein G8599;
 C/Spectra: Saccharomyces cerevisiae
 C/Date: 31-Mar-1990 #sequence, revision 14-Jul-1994 #ext-change 09-Jul-2004
 C/Accession: J00016; S38963; S57708; S64566; S64564; S63924

R.Helms, J.; Ritzel, R.G.; von Borestel, R.C.; Aguilera, A.; Rodicio, R.; Zimmermann,
 Gene 78, 309-321, 1989

A>Title: The phosphofructokinase genes of yeast evolved from two duplication events.

A/Reference number: A91608; MUID:89378757; PMID:2528496

A/Accession: J00016

A/Molecule type: DNA

A/Residues: 1-987 <HEI>

A/Cross-references: UNIPROT:P16861; EMBL:M26943; NID:g172137; PIDN:AA34859.1; PID:g1721

R.Koperechleeger, G.; Baer, J.; Stettinwagen, E.

Eur. J. Biochem. 217, 527-533, 1993

A>Title: Limited proteolysis of yeast phosphofructokinase. Sequence locations of cleavag

A/Reference number: S38963; MUID:94039086; PMID:8223596

A/Accession: S38963

A/Molecule type: protein

A/Residues: 1-6;90-97;197-205;914-921 <KOP>

R.van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.

submitted to the EMBL Data Library, June 1995

A/Description: Sequence analysis of the 43 KB CRM1-YLM9-PET54-SM11-PHO81-YHB4-PFK1 regio

A/Reference number: S57680

A/Accession: S57708

A/Molecule type: DNA

A/Residues: 794-987 <VAN>

A/Cross-references: EMBL:X87941; NID:g886908; PIDN:CAA61193.1; PID:g886937

A/Experimental source: strain S288C

R.Guerreiro, P.; Barreiro, T.; Azevedo, D.; Rodrigues-Pousada, C.

submitted to the Protein Sequence Database, May 1996

A/Reference number: S64565

A/Accession: S64566

A/Molecule type: DNA

A/Residues: 1-987 <GUB>

A/Cross-references: EMBL:Z73025; NID:g1323434; PIDN:CAA97268.1; PID:g1323435; GSPDB:GN00C

A/Experimental source: strain S288C

R.van der Aart, Q.J.M.; Steensma, H.Y.

submitted to the Protein Sequence Database, May 1996

A/Reference number: S64541

A/Accession: S64564

A/Molecule type: DNA

A/Residues: 794-987 <VAN>

A/Cross-references: EMBL:Z73025; GSPDB:GN00007; MIPS:YGR240C

A/Experimental source: strain S288C

R.van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.

Yeast 12, 385-390, 1996

A>Title: Sequence analysis of the 43 kb CRM1-YLM9-PET54-DIE2-SM11-PHO81-YHB4-PFK1 region

A/Reference number: S63896; MUID:96267763; PMID:8701610

A/Accession: S63924

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 794-987 <VAN>

A/Cross-references: EMBL:X87941; NID:g886908; PIDN:CAA61193.1; PID:g886937

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C/Comment: Phosphofructokinase is composed of four alpha chains and four beta chains.

C/Genetics:

A/cene: SGD:PFK1; MIPS:YGR240C

A/Cross-references: SGD:S0003472; MIPS:YGR240C

A/Map position: 7R

C/Superfamily: ATP-dependent phosphofructokinase, eukaryotic type; 6-phosphofructokinase

C/Keywords: ATP; glycolysis; phosphotransferase

F,208-517/Domain: 6-phosphofructokinase 1 homology <6PF>

F,596-888/Domain: 6-phosphofructokinase 1 homology <6PF>

F,325,229/Binding site: AMP, allosteric (Arg) #status predicted

F,356/Active site: Asp (Asp, Gly) #status predicted

F,398,455,488,491,729/Binding site: fructose-6-phosphate (Met, Glu, Lys, His, Arg, G)

F,751/Binding site: citrate, allosteric (Lys) #status predicted

Query Match 10.8%; Score 246.5; DB 1; Length 987;

Best Local Similarity 23.9%; Pred. No. 2.8e-09;

Matches 103; Conservative 68; Mismatches 135; Indels 125; Gaps 19;

QY 31 KKKVAAILNGLAPCLNSALISLERYEIDPSIEIICRGYKGLLDGSPYTAER 90
 DB 204 KKKKIAVMTSGDSPQMAVAVAVR--TGIFGCDVPAVYGBELRGKTK---LK 256
 QY 91 KKA-----GLRFGGSVGNRSVKLTNVKDCVKRGLVKEGDEPOKVAADQLVKDGVDIL 145
 DB 257 KNAEMEDVKGWLB--GGTILGT-----RSMERKREGRQAAGNLISGIDIL 303
 QY 146 HTGGDDTNTAAADL-----AFLARNY-----GLTVIGLPTVNDVFP1 187
 DB 304 VVCGGGSGLT--GADLFRHEWPSLVDELVAEGFTEEVAPYKRLSIVGLVGSIDNDMSGT 362
 QY 188 KQSLGAMTAEGGARYFNNVNAENNANPMLIVHEMGNGCMTL---AATPOEYRKL 243
 DB 363 DSTIGVSLERICEVNDYIDATFASHSRAPFV--EWMGHGCMWLLMAGIATGADY---- 417
 QY 244 DBAEMLPBGLTRSEYVAHVPEMAIDLBABAKRLREVMKQVDCV-----NIFPSGA 298
 DB 418 -----TIPERAVPHGKODELVCORHNSKGRANNTIYAEGA 457
 QY 299 -----GVEAIVAEQKGOEVPDAFGHKL---DAVNPQKMFGEOPOMIAEK 345
 DB 458 LDDQLNPVTANVYKALIEL---GLDTKTYITIGHVQGGTAVAHBWLAT--LQGVDAVK 512
 QY 346 TLVQKSGYFARASASN-----VDMRLIKSCADLAVECAFRBSGVYGHDE 391
 DB 513 AYLE-----FTPTSPPLIGILENKTIIRMLVESVYLKTSVA-----TALENKOF 557
 QY 392 DNGNVTARATIEF 402
 DB 558 DKALISLDRTEF 568

RESULT 11
S32902
6-phosphofructokinase (EC 2.7.1.11) alpha chain - yeast (*Kluyveromyces marxianus* var. 1a
C)Species: *Kluyveromyces marxianus* var. *lactis*, *Candida sphaerica*
C)Date: 08-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C)Accession: S32902; S29602
R)Heinisch, J.; Kitchinath, L.; Liseen, T.; Vogelwang, K.; Hollenberg, C.P.
Mol. Microbiol. 8, 559-570, 1993
A)Title: Molecular genetics of phosphofructokinase in the yeast *Kluyveromyces lactis*.
A)Reference number: S32902; MUID:93316853; PMID:8326866
A)Accession: S32902
A)Molecule type: DNA
A)Residues: 1-992 <HE>
A)Cross-references: UNIPROT:003215; EMBL:Z17315; NID:g2879; PIDN:CAA78963.1; PID:g2880
C)Genetics:
A)Gene: PFK1
C)Superfamily: ATP-dependent phosphofructokinase, eukaryotic type; 6-phosphofructokinase
C)Keywords: ATP, phosphotransferase
F)186-499/Domain: 6-phosphofructokinase 1 homology <6PF1>
F)1574-866/Domain: 6-phosphofructokinase 1 homology <6PF2>

Query Match 10.7%; Score 245.5; DB 2; Length 992;
Best Local Similarity 23.8%; Pred. No. 3.3e-09;
Matches 108; Conservative 74; Mismatches 165; Indels 107; Gaps 19;
QY 30 NKKPKVAITLTAAGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLDGSPVTAEV 89
DB 181 SKKKKIAVMTSGDGSQGMMAVRAVVR--SSIYGCDAVAAYEGEGLVKGDDYLRKMEW 238
QY 90 RKKAAVLQRFQGSVIGNSRVKLTNVKDCVKGRLVEGED--PQKAAVQVLVMDGDIILHT 147
DB 239 KQVRKMLSB-GSTLIGTA-----SKPFRMRKQACSNLIDQGIDALVY 283
QY 148 IGGDDTNTAA-----ADLAAFLARN-----YGLTVIGLPTVNDVPIKOS 190
DB 284 IGGDSSLGADLFRSEWPSLVEIVKDGKTEDEVAVLQNLITVGMVSGIDMDSGTST 343
QY 191 LGAMTAAEQAGARYFMNVVAENNANPRLIVHEVMGNGCWL---TAATAQYRKLDPRA 246
DB 344 IGAVSALERICEMVDYIDATAKSHSRAFAV--EVMGRHCGMLGMSGIAVTAADY----- 395
QY 247 EMLPRLGLTRESYEVAHVAVPEMAI-----DLBAKRLREVDKVCVNIPIVSEGA- 298
DB 396 -----LFIPEAAPHGKMDQLKRCQRRHEKGRNN--TVIAEGGL 436
QY 299 --GVEAIVAEQAK-----GOEVRDAFGHKL--DAVNPQKMFGEQFAQWIGAEKTLV 348
DB 437 DDQLNPITAE-QVKQVLELGLDTKITTGHTGVRGGTAVAHDRMLAT--LQGVDAVKAIL 493
QY 349 ---OKSGYFARASASNVDMRLIKSCADLAVECAFRESGVIGHEDNGVLAIEPFR 404
DB 494 NMTPEPSPILGILNKVIRMPLES-----VKTRQVAAAIKADFPKATISLRDTEFE 548
QY 405 I-----KGGKPFNIDTDMFMSLSEIQP 428
DB 549 LYSNFMSTTVDGSQLPEADRLNIAIVHGA 582

RESULT 12
T10691
probable diphosphate-fructose-6-phosphate 1-phosphotransferase (EC 2.7.1.90) T16118.50 -
N)Alternate names: protein T16118.50
C)Species: *Arabidopsis thaliana* (mouse-ear cress)
C)Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C)Accession: T10691
R)Bayan, M.; Lemnard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, June 1999
A)Reference number: Z16533
A)Accession: T10691
A)Molecule type: DNA
A)Residues: 1-462 <BEV>
A)Cross-references: UNIPROT:Q9M076; EMBL:AL049915; GSPDB:GN00062; ATSP:T16118.50
A)Experimental source: cultivar Columbia; BAC clone T16118

C)Genetics:
A)Gene: ATSP:T16118.50
A)Map position: 4
A)Introns: 35/3; 51/2; 64/3; 85/3; 131/3; 184/3; 202/3; 225/3; 269/1; 354/3; 371/1; 422/3;
C)Superfamily: pyrophosphate-dependent phosphofructokinase, T10108 type; 6-phosphofructo
C)Keywords: phosphotransferase
F)95-403/Domain: 6-phosphofructokinase 1 homology <6PF>

Query Match 10.6%; Score 243; DB 2; Length 462;
Best Local Similarity 25.2%; Pred. No. 1.8e-09;
Matches 105; Conservative 58; Mismatches 178; Indels 76; Gaps 14;
QY 31 KPKKV--AIIATAGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLDGSPVTAEV 88
DB 89 KPSDRACIVTCGGICPELNTVIREIVYVGLHFMGVTEIVGDCGFRFPYSKNTALTP- 147
QY 89 VRKAGVLQRFQGSVIGNSRVKLTNVKDCVKGRLVEGEDPQKAAVQVLVMDGDIILHT 148
DB 148 --KIVSDIHKRGGITLIGSR-----GCHDTSKI--VDNIQDREINQVYII 188
QY 149 GDDDTNTAAADLAAFLARNNGYGLTVIGLPTVNDVPIKOSLGAWTAEGAGARYFMNV 208
DB 189 GDDGTQKGANIYKEIRRRGLKVAAGIPKTIIDIPVIDKSFQDFAVEEQRAINAAH 248
QY 209 AENNANPRLIVHEVMGNGCWLTAAATQAEYRKLDPRAEMLPGLTRESYEVAHVAVE 268
DB 249 VEATSVENGIGIVKLMGRYSGFIA-----MYATLASRDVDCCLIP- 289
QY 269 MAIDLEA-----AKRLREVMKVCVNIPIVSEGAQVEAI---VAEMQAKGQEVPRDA 318
DB 290 SPFLIEGKGYEFLAQLRNGHNV-----IVIAEGAGQDLVABSIHQDASGNKLLNDV 345
QY 319 FGHIKLDVNPQKFG---EQFAQWIGAEKTLVQ--KSGYFARASASNVDMRLIKSCAD 373
DB 346 -----GLWMSLTIKEYFAKHVMYDITLKTIPTYIMRAIPANASNVYSTLLAQ 394
QY 374 LAVECAPRESGVIGHEDNGVLAIEPFR--KGGKPFNIDTDMFMSLSEIQP 428
DB 395 SAVGNMAGYGFVSGLVNGHRY--IPFNRIITERQNKRVITDRMMAR--MLSTNQP 448

RESULT 13
S54978
6-phosphofructokinase (EC 2.7.1.11), pyrophosphate-dependent - *Neegleria fowleri*
N)Alternate names: pyrophosphate-dependent phosphofructo-1-kinase
C)Species: *Neegleria fowleri*
C)Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004
C)Accession: S54978; S71070
R)Wesberg, K.L.; Skolnick, S.; Xu, J.; Marciano-Cabral, F.; Kemp, R.G.
Biochem. J. 307, 143-149, 1995
A)Title: Cloning, sequencing and expression of the pyrophosphate-dependent phosphofructo-
A)Reference number: S54978; MUID:95234028; PMID:7717968
A)Accession: S54978
A)Molecule type: mRNA
A)Residues: 1-437 <RES>
A)Cross-references: UNIPROT:Q27705; EMBL:U11733
R)Kemp, R.G.
submitted to the EMBL Data Library, July 1994
A)Reference number: S71070
A)Accession: S71070
A)Molecule type: mRNA
A)Residues: 1-141, 'VTL', 145-437 <KEM>
A)Cross-references: EMBL:U11733; NID:9511230; PIDN:AAA65791.1; PID:9511231
C)Superfamily: pyrophosphate-dependent phosphofructokinase, T10289 type; 6-phosphofructo
C)Keywords: phosphotransferase
F)20-352/Domain: 6-phosphofructokinase 1 homology <6PF>

Query Match 10.6%; Score 241.5; DB 2; Length 437;
Best Local Similarity 27.6%; Pred. No. 2.1e-09;
Matches 120; Conservative 65; Mismatches 171; Indels 79; Gaps 24;
QY 27 TLNKKPKV---AIIATAGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLDGDS 82

Db 9 TTIVTPKAVPTLVGVGGPAPGIVGAVTIE---AATNGRVULGFIEGPNLLODD 65
Qy 83 YPVTAERKKAQVLOFGSGVIGNSRVKLTNNVCYKGLVKEGEDPOKVAADQLVKDGV 142
Db 66 SKIVELTIDSVSRHIEGGSIILKTSRANPT-----KKQEDLOQVY-VKQLOKFNV 113
Qy 143 DLTHTIGSDPTTAAADLAFLARNNYGLTIGLPTVNDVFPKQSL---GAMTAAQ 199
Db 114 SLVYIIGSD--TAFSSMVAKAAANN-ELHVCVPTINDL-PLFYGIPTFGYEYARBF 169
Qy 200 GARYFNVAENNANPRMLIYHEVWGRNGWLT---AATAQERKLDIAEMLPGLR 256
Db 170 GANVVRNLTMTDASTASRYIV-VAMGRQAGHLALGKSAQSHLILPPE-EFLP-----TT 223
Qy 257 ESVEVAVF--VPEMAIDLEAF-AKRLREVMKVCNIFVSEGAVEAIVAE--QAKQ 312
Db 224 DSTEPEVTSRICDM---IEASIIKRL--YTSKKDHGVILAEGLLEYNSTDELKAFSS 278
Qy 313 EYPRDAFGHIKLDVAVNGKWFGEQFAQMT---GAETLVQKS-GYPARASASNVDMRLI 368
Db 279 SKYDADHIMLAELDFGLVDEMEREMRRRLKIAFTEKNLGYELRCAPVAFPREYT 338
Qy 369 KSCADAAVCAFRRESGVIGHEDNGNVLAIEFPRIKGG--PRNIDDMFSMLSEIG 426
Db 339 RULGNAVAVYLLANGNGAL-----ITVQGVKVPVLSFD-----DLK 374
Qy 427 QPKGK-----VEVS 436
Db 375 DPRTGKTRTROYDVS 389

RESULT 14

AH0942
6-phosphofructokinase [imported] - Salmonella enterica subsp. enterica serovar Typhi (serotype 6)
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-May-2004
C:Accession: AH0942
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moul, S.; O'Gaora, P.
Nucleotide: 848-852, 2001
A:Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi: AB0502; PMID:11677608
A:Reference number: AB0502; PMID:11677608
A:Accession: AH0942
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-320 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD09562.1; PID:gl6504676; GSPDB:GN00176
C:Genetic: 6
A:Gene: STY3809
C:Superfamily: ATP-dependent phosphofructokinase, prokaryotic type; 6-phosphofructokinase

Query Match 10.5%; Score 240.5; DB 2; Length 320;
Best Local Similarity 24.8%; Pred. No. 1.6e-09;

Matches 102; Conservative 59; Mismatches 122; Indels 129; Gaps 18;

Qy 33 KVAIITAGLAPCLNSAIGSLIERTEIDPSIEITCYGKYGK-----LLGDSFPVTA 87
Db 3 KRIGVLTSGDAPGMMAAIGVVR--AALTEGLEWGIYDGLYEDRMVOLDRYSVD 60
Qy 88 EVRKKAQVLOFGSGVIGNSRVKLTNNVCYKGLVKEGEDPOKVAADQLVKDGVILHT 147
Db 61 MNR-----GGTILGSR--FPFPRD-----ENRAVAILENKKRGIHALV 100
Qy 148 IGGDDTNTAAADLAFLARNNYGLTIGLPTVNDVFPKQSLGAMTAAEQARYFNV 207
Db 101 IGGDGSYMAKRL-----TEMGFCIGLPGTIDNDIKGDYIGFTALGVTAIDRL 154
Qy 208 VAEENANPRMLIYHEVWGRNGWLTAAQERKLDIAEMLPGLTRESYEVHAFV 265
Db 155 RUTSSHORISIV-EVMGRYCGDLTAA-----IAGGCEFIIVPEVERNR----- 200

Qy 266 VPEMAIDLEAFK-----RLREVMKVCNIFVSEGAVF--AIVAEQAK 310
Db 201 -----DLVAEIKAGIAKKAIVAITHEMCDVDLAFIEKETERTRAATVIGHIQG 254
Qy 311 GQVPRDAFGHIKLDVAVNGKWFGEQFAQMTGAETLVQKSGYPARASASNVDMRLIS 370
Db 255 GSPVPD-----RLASRMKAYV----- 272
Qy 371 CADLAVE-----CAFRRESGVIGHEDNGNVLAIEFPRIKGGKPPNIDTW 417
Db 273 -IDLLEHGRCVGIQNEQLVHND-----IIDALENMK-----RPFK--SDW 312

RESULT 15

KIECPA

6-phosphofructokinase (EC 2.7.1.11) 1 - Escherichia coli (strain K-12)
N:Alternate names: phosphofructokinase 1, isozyme 1; phosphohexokinase, isozyme 1
C:Species: Escherichia coli
C:date: 31-Dec-1988 #sequence_revision 10-Oct-1997 #text_change 09-Jul-2004
C:Accession: G65197; A25206; S40859
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coi, A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; PMID:97426617; PMID:9278503
A:Accession: G65197
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-320 <BLAT>
A:Cross-references: UNIPROT: P06998; GB:AE000466; GB:U00096; NID:g2367328; PIDN:AACT6898.J
A:Experimental source: strain K-12, substrain MG1655
R:Heilling, H.W.; Evans, P.R.
Eur. J. Biochem. 149, 363-373, 1985
A:Title: Nucleotide sequence and high-level expression of the major Escherichia coli phosphofructokinase gene
A:Reference number: A91144; PMID:85203917; PMID:3158524
A:Accession: A25206
A:Molecule type: DNA
A:Residues: 1-73, 'C', '75-102, 'DG', '105-162, 'P', '164-316, 'E', '318, 'M', '320 <HEL>
A:Cross-references: GB:X02519; NID:G42365; PIDN:CAA26356.1; PID:g42366
A:Note: this sequence has since been corrected
R:Evans, P.R.
submitted to the EMBL Data Library, October 1986
A:Reference number: A94501
A:Content: annotation; corrections
R:Plunkett III, G.; Burland, V.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 21, 3391-3398, 1993
A:Title: Analysis of the Escherichia coli genome. III. DNA sequence of the region from 8, A:Reference number: S40802; PMID:93347969; PMID:8346018
A:Accession: S40859
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-320 <PUB>
A:Cross-references: EMBL: L19201; NID:g304961; PIDN:AA03048.1; PID:g305019
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, October 1993
C:Comment: The active enzyme catalyzes the key control step of glycolysis, the phosphorylation of ADP and inhibited by phosphoenolpyruvate.
Y:ADP and inhibited by phosphoenolpyruvate.
C:Comment: In E. coli this enzyme is responsible for 90% of the phosphofructokinase activity
Y:ADP and inhibited by phosphoenolpyruvate; there is also no immunological cross-reactivity.
C:Genetic: 6
A:Gene: pfkA
A:Map position: 88 min
C:Superfamily: ATP-dependent phosphofructokinase, prokaryotic type; 6-phosphofructokinase
C:Keyword: allosteric regulation; ATP; glycolysis; homotetramer; phosphotransferase
P:3-219/Domain: 6-phosphofructokinase 1 homology <6PPI>

Query Match 10.4%; Score 238.5; DB 1; Length 320;
Best Local Similarity 24.9%; Pred. No. 2.2e-09;

Matches 100; Conservative 65; Mismatches 132; Indels 105; Gaps 18;

Qy 33 KVAIITAGLAPCLNSAIGSLIERTEIDPSIEITCYGKYGK-----LLGDSFPVTA 87
Db 3 KRIGVLTSGDAPGMMAAIGVVR--AALTEGLEWGIYDGLYEDRMVOLDRYSVD 60

```
QY      88 EYRKKAQVUQRFGSVIGNSRYKLTIVKDCVKGGLYKEGEDPQKVAADQLVKQVDILHT 147
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61 MINR-----GGTFLGSR--FPEFRD-----ENIRAVAIENLKRRGIDALVV 100
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      148 IGGDDTNTAAADLAAFLAARNNYGLTVIGPKTVNDVFPPIKOSLGAWTAEOGARFEMV 207
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      101 IGGDGSYMGAMRL-----TEMGFPCIGLPGTIIDNDIKGTDYTTIGFTALSTVVEAIDRL 154
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      208 VAENNANPRLIVHEVWGNCGMLTAATAOEYRKLDRAEW--LPELGLTRESYEVAVF 265
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      155 RDTSSHQRI SYV-EVMGRYCGDLTLAA-----IAGGCEFYVPEVEFSRED-----L 202
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      266 VPEMAIDLEAEAKR-----LREVMKVDCVNI FVSEGAVE--AI VAE MOAKQOEVP RD 317
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      203 VNEIRKAGI-AKGRKAIVAITEHMCVDDELAHFIETKGTRETRATVTLGHIQRGGSPVPYD 261
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      318 AFGHITKLDAVNPGRKMGEGQFAQMIGAEKTLVOKSGYFARASASNVDMRLIKSCADLAVE 377
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      262 RI-----LASRMGAYAIIDL LLAGYGR----- 283
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      378 CAFRRRESGVIGHDEDNQNVLRALIEPPRIKGGKPFNIDTDFN 419
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      284 CVGIQNEQLVHHD-----IIDAIBNMK---RPFK--GDWLD 314
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: February 24, 2005, 14:23:20
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2005, 08:27:31 ; Search time 5943 Seconds
(without alignments)
3563.003 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTWPHYHTRADIRCFHNF.....FNSMLSEIQPKGKGVESH 437

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n model -DEV=xlh
-O=/gen2.1/USPTO/seqol/US10701200/runat_25022005_102832_10310/app_query.fasta_1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10701200 @CCN 1 1 3731 @runat_25022005_102832_10310 -NCPUS=6 -ICPU=3
-NO MAP -LARGEDEUTER -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*

1: gb_ba:*\n2: gb_hcg:*\n3: gb_in:*\n4: gb_om:*\n5: gb_ov:*\n6: gb_pat:*\n7: gb_ph:*\n8: gb_pl:*\n9: gb_pr:*\n10: gb_ro:*\n11: gb_scs:*\n12: gb_un:*\n13: gb_vl:*\n14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	2284	100.0	1311	6 AR308857 Sequence
2	2284	100.0	1311	6 AR473868 Sequence
3	2284	100.0	1311	6 AR565849 Sequence
4	2284	100.0	1311	6 AX394958 Sequence

5	2284	100.0	1311	6 AX398319 Sequence
6	2284	100.0	1311	6 AX398428 Sequence
7	1531.5	67.1	295650	1 BX294151 Pirelinula
8	1453	63.6	1427	3 AP246209 Mastigamo
9	1444.5	63.2	8530	6 CQ363740 Sequence
10	1434.5	62.8	11000	1 AE017283_11
11	1314	57.5	1400	1 PRSPFK
12	1302	57.0	294800	1 SWS591789
13	1292.5	56.6	11715	1 AE008127
14	1292.5	56.6	11807	1 AE009161
15	311	13.6	1549	1 SCU51728
16	311	13.6	321250	1 SC0939111
17	305	13.4	1233	1 AF268276
18	298.5	13.1	2302	1 AMU31277
19	298	13.0	299925	1 AP005045
20	284.5	12.5	299425	1 AP005049
21	281.5	12.3	1458	8 AY040055
22	281.5	12.3	2087	8 AF360207
23	281.5	12.3	298050	1 AP003189
24	281.5	12.3	339650	1 SC0939108
25	279.5	12.2	110000	1 AE016822_07
26	279	12.2	141027	8 AC132486
27	279	12.2	163670	8 AC130730
28	275	12.0	110000	1 CP000013_7
29	274.5	12.0	2490	1 AF336847
30	272.5	11.9	110000	2 AP006495_4
31	271.5	11.9	110000	1 AE017180_22
32	271.5	11.9	300100	1 SC0939123
33	270	11.8	1778	6 AR050093
34	270	11.8	11813	1 AE001172
35	269.5	11.8	281450	1 AP005032
36	267.5	11.7	300205	1 AE015944
37	266	11.6	1467	8 BT015466
38	266	11.6	1470	8 AY090267
39	266	11.6	1990	8 AY049245
40	264.5	11.6	1624	6 AR050094
41	263.5	11.5	1327	3 EHPPIPEK
42	263.5	11.5	2256	6 AX654108
43	262	11.5	1978	6 AR050092
44	262	11.5	15158	1 AE001705
45	261.5	11.4	1490	6 AX654128

ALIGNMENTS

RESULT 1
AR308857 LOCUS AR308857 Sequence 15 from patent US 6555353.
DEFINITION AR308857
ACCESSION AR308857
VERSION AR308857.1 GI:31700588
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1311)
AUTHORS Koffas, M., Norton, K.C., Odom, J.W., and Ye, R.W.
TITLES Methanotrophic carbon metabolism pathway genes and enzymes
JOURNAL Patent: US 6555353-A 15 29-APR-2003;
FEATURES
source 1..1311
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2.12e-161
Score: 2284.00 Length: 1311
Percent Similarity: 100.00% Matches: 437
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-701-200-6 (1-437) x AR308857 (1-1311)

QY 1 AspValValThrTrpProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhe 20
 Db 1 GATGGTGCACATGGCCCTATCATCACTTAAGCGCTGATTCGATTTGTCATTTGTTTTT 60
 QY 21 LeuAnpPheAnpPheTyrThrLeuMetAnLysProLysLysValAlaIleLeuThrAla 40
 Db 61 CTTAACTTTAACTTCTACACGCTCATGAACCAACCTAAAGAGTGCATCTGACAGCA 120
 QY 41 GlyGlyLeuAlaProCysLeuAnSerAlaIleGlySerLeuIleGluArgTyrThrGlu 60
 Db 121 GCGCGCTTGCGCCCTTTGTTGATTCGCAATCCGATCGATTCGACAGCTTATACCAA 180
 QY 61 IleAppProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGlyLeuLeuLeuGly 80
 Db 181 ATCGATCTTACGATAGAAATCATTTGCTATCGCGGGGCTTATAAGGCGCTGTTGCTGG 240
 QY 81 AppSerTyrProValThrAlaGluValArgLysLysValAlaIleAspGluLeuValLysAsp 100
 Db 241 GATTTCTTATCCAGTAAAGCGCCGAGAGCGTAAAGCGGGGTCTTCGCAAGCTTTGGC 300
 QY 101 GlySerValIleGlyLysnSerArgValLysLeuThrAnValLysAspCysValLysArg 120
 Db 301 GGTTCGTGATCGGCACACGCGCGCTCAAAATTCGACCAATGTCGAAAGCTGCTGAAGC 360
 QY 121 GlyLeuValLysGlyGlyGluAspProGluLysValAlaIleAspGluLeuValLysAsp 140
 Db 361 GGTTCGTCAAGAGGGGTGAAGATCCGCAGAAAGTCCGCGCTGATCAATTTGGTTAAGAT 420
 QY 141 GlyValAspIleLeuHisThrIleGlyLysAspProThrAnThrAlaIleAspLeu 160
 Db 421 GGTGTCGATATCTGCACACCATCGCGCGGATGATACCAATACGCGACGACGGAATTTG 480
 QY 161 AlaAlaPheLeuAlaArgAnAsnTyrGlyLeuThrValIleGlyLeuProLysThrVal 180
 Db 481 CGAGCATTCCTGCGCAAGAAATTAATTCGACCTGACCGCTATGCTTACTTAAACCGTC 540
 QY 181 AppAnpAspValPheProIleLysGlnSerLeuGlyValaTrpThrAlaAlaGluGlnGly 200
 Db 541 GATTAACGACGATATTCCTGATCAAGCAATCATAGGTGCTTGGACCTCGCCGACAGAGGC 600
 QY 201 AlaArgTyrPheMetAnValValAlaGluAnAnAlaAnProArgMetLeuIleVal 220
 Db 601 GCGCGTATTTATCATGAACGTGGTGGCGCAAAACAACGCGCAACCGCATGTGATCGTA 660
 QY 221 HisGluValMetGlyArgAnCysGlyTyrPleuThrAlaAlaThrAlaGlnGluTyrArg 240
 Db 661 CACGAAGTATGGCGCGTAACTGCGGCTGGCTGACCGCTGCAACCGCGAGAGAAATTCGC 720
 QY 241 LysLeuLeuAspArgAlaGluTrpLeuProGluLeuGlyLeuThrArgGluSerTyrGlu 260
 Db 721 AAATTAACCTGACCGTGGCCGAGTGGTGGCGGAATTTGGTTGACTGCTGTAATCTTATGA 780
 QY 261 ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluAlaLysArgLeu 280
 Db 781 GTGCGCGCGGTATTCGTTCCGAAATGGGATCGATCGATCGAAGCCGAAGCCGAAGCGCTG 840
 QY 281 ArgGluValMetAspLysValAspCysValAnIlePheValSerGluGlyValaGlyVal 300
 Db 841 CCGCAAGTATGATGCAAACTCGATTCGCTCAACATCTTCGTTCCGAAGTATCCGCGCTC 900
 QY 301 GluAlaIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
 Db 901 GAAGCTATGTCGCGAAATGCAGGCGCAAGGCGAGAAAGTCCGCGCGCATCGTTCCGCG 960
 QY 321 HisIleLysLeuAspAlaValAnpProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
 Db 961 CACATCAAACTGATGCGGTCAACCTCGGTAAATGGTTCCGGACCAATTCGCGAGATG 1020
 QY 341 IleGlyAlaGluLysThrLeuValGluLysSerGlyTyrPheAlaArgAlaSerAlaSer 360
 Db 1021 ATAGCGCGGAGAAACCTGATCAAAATCGGATACTTCGCGCGTCTTCGCTTC 1080

QY 361 AnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
 Db 1081 AACGTTGACGACATGCGTTTATCATCAATCGTCCGCGCTGGCGGTGAGAGCCGCTTC 1140
 QY 381 ArgArgGluSerGlyValIleGlyHisAspGluAspAnGlyAnValLeuArgAlaIle 400
 Db 1141 GCGCGCGGTCTGGGGTATCGTCCACAGCAAGCAACGCGCAACGTGTGCTGCTGATC 1200
 QY 401 GluPheProArgIleLysGlyGlyLysProPheAnIleAspThrAspTrpPheAnSer 420
 Db 1201 GAGTTCCGCCCATCAAGGGGGGCAACCGTTCAATATCGACACCGCATGTTCAATAGC 1260
 QY 421 MetLeuSerGluIleGlyGlnProLysGlyGlyLysValGluValSerHis 437
 Db 1261 ATGTTGACGAAATCGCGCACGCTTAAGCGGCTTAAGTCGAAGTCACGAC 1311
 RESULT 2
 AR473868
 LOCUS AR473868 1311 bp DNA linear PAT 20-FEB-2004
 DEFINITION Sequence 5 from patent US 6689601.
 ACCESSION AR473868
 VERSION AR473868.1 GI:42712367
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1311)
 AUTHORS Koffas, M., Odom, J.M., and Schenzle, A.
 TITLE High growth methanotropic bacterial strain
 JOURNAL Patent: US 6689601-A 5 10-FEB-2004;
 FEATURES
 source
 1. 1311
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,12e-161 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-10-701-200-6 (1-437) x AR473868 (1-1311)
 QY 1 AspValValThrTrpProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhe 20
 Db 1 GATGGTGCACATGGCCCTATCATCACTTAAGCGCTGATTCGATTTGTCATTTGTTTTT 60
 QY 21 LeuAnpPheAnpPheTyrThrLeuMetAnLysProLysLysValAlaIleLeuThrAla 40
 Db 61 CTTAACTTTAACTTCTACACGCTCATGAACCAACCTAAAGAGTGCATCTGACAGCA 120
 QY 41 GlyGlyLeuAlaProCysLeuAnSerAlaIleGlySerLeuIleGluArgTyrThrGlu 60
 Db 121 GCGCGCTTGCGCCCTTTGTTGATTCGCAATCCGATCGATTCGACAGCTTATACCAA 180
 QY 61 IleAppProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGlyLeuLeuLeuGly 80
 Db 181 ATCGATCTTACGATAGAAATCATTTGCTATCGCGGGGCTTATAAGGCGCTGTTGCTGG 240
 QY 81 AppSerTyrProValThrAlaGluValArgLysLysValAlaIleAspGluLeuValLysAsp 100
 Db 241 GATTTCTTATCCAGTAAAGCGCCGAGAGCGTAAAGCGGGGTCTTCGCAAGCTTTGGC 300
 QY 101 GlySerValIleGlyLysnSerArgValLysLeuThrAnValLysAspCysValLysArg 120
 Db 301 GGTTCGTGATCGGCACACGCGCGCTCAAAATTCGACCAATGTCGAAAGCTGCTGAAGC 360
 QY 121 GlyLeuValLysGlyGlyGluAspProGluLysValAlaIleAspGluLeuValLysAsp 140

```

Db 361 GGTTCGTCAGAAAGGGGTGAAGATCCGCAAAAAGTCGGCGCTGATCAATTGGTTAAGAT 420
Qy 141 G1yVala1Asp1IleuH1sthr1IleG1y1AspAspThrAsnThr1Ala1AlaAspLeu 160
Db 421 GGTTGTCATATTCGACACACATCCGGCGCGATGATACCAATACCGCAGACGGGATTTG 480
Qy 161 A1a1a1PheLeu1AlaArgAsnAntyrg1yLeuThrVal1IleG1yLeuProlysthrVal 180
Db 481 GCAGCATTCCTGGCCGAGAATAATTAACGACGACGTCATGTTTACTTAACCAACCGTC 540
Qy 181 AAPAsnAspValPheProI1elysg1nserLeuG1yAla1TrpThr1Ala1AlaG1ng1y 200
Db 541 GATTAACGACGTATTCGATCAAGCAATACATAGGCTTGATCCGCCCGAGCAAGGC 600
Qy 201 A1aArgTyPheMetAsnVal1AlaG1uAsnAsn1AsnProArgMetLeu1IleVal 220
Db 601 GCGCGTTATTTTCATGAACGTGGTGGCCGAAACACGCCAACCCACGATGCTGATGTA 660
Qy 221 H1sg1uVal1MetG1yArgAsnCy8g1yTrpLeuThr1Ala1AlaThr1AlaG1ng1yArg 240
Db 661 CACGAAGTATGGCGCGTAATCTCGGCTGACCGCTGACCGGCAACCGCAGAAATTCG 720
Qy 241 LysLeuLeuAspArg1AlaG1uTrpLeuProG1uLeuG1yLeuThrArg1G1uSerTyrg1u 260
Db 721 AAATTACTGACCGTCCGAGTGGTTCGCGAAATGGGATTGACTCGTGAATCTTATGAA 780
Qy 261 Val1H1sAlaVal1PheVal1ProG1uMetAla1IleAspLeuG1uAlaG1uAlaValArgLeu 280
Db 781 GTGCACGCGTATTCCTTCGGAATGGGATGATGATCGATCGAACCCAGCCAGCCCGCTG 840
Qy 281 ArgG1uVal1MetAsp1yAspVal1AspCy8Val1Asn1IlePheVal1SerG1uG1yAlaG1yVal 300
Db 841 CCGGAAGTATGAGCAAAAGTCATCGATCGATCAATCTTCGTTCCGAAAGTTCGCGCGTC 900
Qy 301 G1uVal1IleVal1AlaG1uMetG1nAla1yG1yG1uG1uVal1ProArgAsp1AlaPheG1y 320
Db 901 GAAGCATTCGTCGCGAAATGACAGGCCAAAGCCAAAGTGGCCGCGATGGTTCGCG 960
Qy 321 H1s1IleLysLeuAsp1AlaVal1AsnProG1yLysTrpPheG1yG1uG1n1PheAlaG1Met 340
Db 961 CACATTAACCTGATCGGTCACCTGTTAATGTTTCGCGACGAAATTCGCGCAGATG 1020
Qy 341 I1eG1yAlaG1uLysThrLeuVal1G1uLysSerG1yTyPheAlaArg1Asn1Asp 360
Db 1021 ATAGCGCGGAAAAAACCTGGTACAAAATCGGAAATCTTCGCCCGTCTTCGCTTCC 1080
Qy 361 AsnVal1AspAspMetArgLeu1IleLysSerCy8AlaAspLeu1AlaG1uCy8AlaPhe 380
Db 1081 AACGTTGACGACATGCGTTTGATCAATCGTCGCCGACCTTGGCGGTGCGAGTCCGCGTTC 1140
Qy 381 ArgArgG1uSerG1yVal1IleG1yH1sAspG1uAspAsnG1yAsnVal1LeuArg1Ala 400
Db 1141 CCGCGGAGTCTGGCGTGAATCGGTCACGAGAAACACGCGCAACGTTGCGTGGATC 1200
Qy 401 G1uPheProArg1IleLysG1yG1yLysProPheAsn1IleAspThrAspTrpPheAsnSer 420
Db 1201 GAGTTTCCCGCATCAAGGCGGCAACGCTTCAATATGACACGCACTGCTTCAATAGC 1260
Qy 421 MetLeuSerG1uIleG1yG1nProLysG1yLysVal1G1uVal1SerHis 437
Db 1261 ATGTTGAGCGAAATCGCGCAGCTTAAGGCGGTAAAGTCGAAAGTCAGCCAC 1311

```

```

RESULT 3
ARS56849 1311 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 15 from patent US 6767744.
ACCESSION ARS56849
VERSION ARS56849.1 GI:53981913
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1311)

```

```

AUTHORS Koffas, M., Norton, K.C., Odum, J.M. and Ye, R.W.
TITLE Methanotrophic carbon metabolism pathway genes and enzymes
JOURNAL Patent: US 6767744-A 15 27-JUL-2004;
FEATURES Location/Qualifiers
source 1..1311
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 2,126-161 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-701-200-6 (1-437) x ARS56849 (1-1311)
Qy 1 AspVal1Ala1ThrTrpProTyrrHisLeuThr1AlaAsp1IleArgPheCysHisTrpPhe 20
Db 1 GATGTGTCACATGGCCCTATCATCTTAACGGCTGATATTGATTTGTCACTTGGTTTTT 60
Qy 21 LeuAsnPheAsnPheTyrrThrLeuMetAsnLysProLysLysVal1Ala1IleLeuThrAla 40
Db 61 CTTAACCTTAACCTTACACAGCTCATGAACAAACCTAAAAAAGTTGCAATACTGACAGCA 120
Qy 41 G1yG1yLysLeu1AlaProCysLeuAsnSerAla1IleG1ySerLeu1IleG1uArgTyrrThrG1u 60
Db 121 GCGGCTTGCGCGCTTGTGAATTCGCAATCGGTGATGATTCGAACGTTATACCGAA 180
Qy 61 I1eAspProSer1IleG1uIle1IleCysTyrrArgG1yG1yTyrrLysG1yLeuLeuG1y 80
Db 181 ATGCATCTTAGCATTAATAATCATTTGCTATCCGGCGGTTATTAAGCCCTTGCTCGGC 240
Qy 81 AspSerTyrrProVal1Thr1AlaG1uVal1ArgLysLysAlaG1yVal1LeuG1nArgPheG1y 100
Db 241 GATTTCTTATCCAGTAACGCCCGCAAGTCGTAATAAAGCGCGGTGTTCTGCAACGTTTGGC 300
Qy 101 G1ySerVal1IleG1yAsnSerArgVal1LysLeuThrAsnVal1LysAspCy8Val1LysArg 120
Db 301 GGTTCGTATCGGCAACAGCGCGTCAAAATGACCAATGTCAAAAGCTGCGGAAACGC 360
Qy 121 G1yLeuVal1IleG1yG1uLeuAspProG1nLysVal1Ala1AspG1nLeuVal1LysAsp 140
Db 361 GGTTCGTCAAGAGGCTGAAGATCCGCAAAAAGTCCGCGCTGATCATATTGTTAAGAT 420
Qy 141 G1yVal1Asp1IleuH1sthr1IleG1y1AspAspThrAsnThr1Ala1AlaAspLeu 160
Db 421 GGTCGATATTCGACACACATCGCGCGATGATACCAATACCGCAGACGCGGATTTG 480
Qy 161 A1a1a1PheLeu1AlaArgAsnAntyrg1yLeuThrVal1IleG1yLeuProlysthrVal 180
Db 481 GCAGCATTCCTGGCCGAGAATAATTAACGACGACGTCATGTTTACTTAACCAACCGTC 540
Qy 181 AAPAsnAspValPheProI1elysg1nserLeuG1yAla1TrpThr1Ala1AlaG1ng1y 200
Db 541 GATTAACGACGTATTCGATCAAGCAATACATAGGCTTGATCCGCCCGAGCAAGGC 600
Qy 201 A1aArgTyPheMetAsnVal1AlaG1uAsnAsn1AsnProArgMetLeu1IleVal 220
Db 601 GCGCGTTATTTTCATGAACGTGGTGGCCGAAACACGCCAACCCACGATGCTGATGTA 660
Qy 221 H1sg1uVal1MetG1yArgAsnCy8g1yTrpLeuThr1Ala1AlaThr1AlaG1ng1yArg 240
Db 661 CACGAAGTATGGCGCGTAATCTCGGCTGACCGCTGACCGGCAACCGCAGAAATTCG 720
Qy 241 LysLeuLeuAspArg1AlaG1uTrpLeuProG1uLeuG1yLeuThrArg1G1uSerTyrg1u 260
Db 721 AAATTACTGACCGTCCGAGTGGTTCGCGAAATGGGATTGACTCGTGAATCTTATGAA 780
Qy 261 Val1H1sAlaVal1PheVal1ProG1uMetAla1IleAspLeuG1uAlaG1uAlaValArgLeu 280

```

Db 781 GTGACGGGTTTGTTCCTCGGAAATGGCGATCGACTTGGAAGCCGAAAGCCGCTG 840
 Qy 281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluGlyValIleGlyVal 300
 Db 841 CGCGAAGGATGATGACAAAGTCGATTGCGTCAACATCTTGTTCCGAAGTCCGGCGTC 900
 Qy 301 GGuAaIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgPheAlaPheGly 320
 Db 901 GAAGCTATCGTCGCGAAATGCGAGCCAAAGGCGAGAAAGTCCGCGCATCTGTTCCGC 960
 Qy 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
 Db 961 CACATCAAACTGATGACCGGTCAACCTCGTAAATGCTTCGCGAGCAATTCGCGCATG 1020
 Qy 341 IleGlyAlaGluLysThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSer 360
 Db 1021 ATAGCGCGGAAAAAACCTCGGTACAAAATCGGGTACTTCGCGCGTCTGCTTC 1080
 Qy 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
 Db 1081 AACGTTGACGACATCGCTTGTATCAAAATCGTCGCGCATTCGCGATCGAGCGCGTTC 1140
 Qy 381 ArgArgGluSerGlyValIleIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle 400
 Db 1141 CGCCCGAGTCTCGCGCTGATCGGTCAACGAAAGACACGCAACGTCGTCGTCGATC 1200
 Qy 401 GluPheProArgIleLysGlyGlyLysProPheAsnIleAspThrAspTrpPheAsnSer 420
 Db 1201 GAGTTTCGCGCGCATGAAGGGCGCAACCGTTCAATATGACACCGACTGTTCAATAGC 1260
 Qy 421 MetLeuSerGluIleGlyGlnProLysGlyGlyLysValGluValSerHis 437
 Db 1261 ATGTTGACGGAATGCGCCAGCCTAAAGCGGTAAAGTCGAAGTCAGCCAC 1311

RESULT 4
 AX394958 1311 bp DNA linear PAT 18-MAY-2002
 LOCUS AX394958 1 from Patent WO0218617.
 DEFINITION Sequence 1 from Patent WO0218617.
 ACCESSION AX394958
 VERSION AX394958.1 GI:21066031
 KEYWORDS
 SOURCE
 ORGANISM
 Methylobionas sp.
 Methylobionas sp.
 Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
 Methylococcaceae; Methylobionas.

REFERENCE
 AUTHORS 1
 TITLE Brzostowicz, P.C., Cheng, Q., Dicosimo, D.J., Koffas, M., Miller, E.S.,
 Odum, J.M., Plicataggio, S.K. and Rouviere, P.B.
 Carotenoid production from a single carbon substrate
 Patent: WO 0218617-A 1 07-MAR-2002;
 JOURNAL E.I. DUPONT DE NEMOURS AND COMPANY (US)
 FEATURES
 source Location/Qualifiers

1. 1311
 /organism="Methylobionas sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:418"
 /note="16a"

ORIGIN

Alignment Scores:
 Pred. No.: 2,126-161 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-701-200-6 (1-437) x AX394958 (1-1311)

Qy 1 AspValAlaThrTrpProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhePhe 20
 Db 1 GATGTGTCACATGCGCCCTATCACTTAACGCTGATATTCGATTGTCATTTGCTTTT 60

Qy 21 LeuAsnPheAsnPheTyrThrLeuMetAsnLysProLysValAlaIleLeuThrAla 40
 Db 61 CTTAACTTTAACTTCTACACGCTCATGAAACAACTTAACAAAGTTGCATATCTGACAGA 120
 Qy 41 GlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGlu 60
 Db 121 GCGGCTTGCGGCTTGTTGTTGAAATTCGCGAATCGGTAGTTGATGMACTTAACCGAA 180
 Qy 61 IleAspProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGlyLeuLeuLeuGly 80
 Db 181 ATCATCTCTAGATATGAATATCATTTGTCATCGCGCGGTATTAAGCCCTGTTGCTGGGC 240
 Qy 81 AspSerTyrProValThrAlaGluValArgLysLysAlaGlyValLeuGlnArgPheGly 100
 Db 241 GATCTTATTCAGTAAACCGCGAAGTGGTAAAGGGGGGTCTCGCAAGGTTTGCC 300
 Qy 101 GlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
 Db 301 GGTTCGTGATCGGCAACAGCCGCTCAAAATTGACCAATGTCMAAGATCGCTGAAGCC 360
 Qy 121 GlyLeuValLysGluGlyLysAspProGlnLysValAlaAlaAspGlnLeuValLysAsp 140
 Db 361 GGTTCGTCAAGAGGGGTGAAGATCCGCAAAAAGTCGCGCTGATCAATTGGTTAAGAT 420
 Qy 141 GlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAlaAlaAspLeu 160
 Db 421 GGTCGTGATATTCGCACACCATCGCGCGATGATACCAATACGGCAGCGGATTTG 480
 Qy 161 AlaAlaPheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLysThrVal 180
 Db 481 GCAGATTCCTGCGCCAGAAATTAATTACGACCTGACCTGTTGTTACTTAAACCGTC 540
 Qy 181 AspAsnAspValPheProIleLysGlnSerLeuGlyAlaTrpThrAlaAlaGluGlnGly 200
 Db 541 GATAACGACGATATTCGATCAAGCAATCATAGGTGCTTGACATGCGCGCAGACAGGC 600
 Qy 201 AlaArgTyrPheMetAsnValValAlaGluAsnAsnAlaAsnProArgMetLeuIleVal 220
 Db 601 GCGCGTTATTTCAATGAAAGGTGGTGGCCGAAACCAACGCGCAACCTGCTGATCGTA 660
 Qy 221 HisGluValMetGlyArgAsnCysGlyTyrPheLeuThrAlaAlaThrAlaGlnGluTyrArg 240
 Db 661 CACGAAAGTATGGGCCCTGTAATCTCGCGCTGCTGACCGCTGCAACCGCGCAGAAATATCC 720
 Qy 241 LysLeuLeuAspArgAlaGluTyrPheLeuProGluLeuGlyLeuThrArgGluSerTyrGlu 260
 Db 721 AAATTACTGACACCGCGGAGGTTGCCGGAATGGGTTGACTCGTGAATCTTATGAA 780
 Qy 261 ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluAlaLysArgLeu 280
 Db 781 GTGACGCGGATATTCGTTCCGAAATGCGATCGACTGGGAAGCCGAAGCCAAAGGCCCTG 840
 Qy 281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluGlyValIleGlyVal 300
 Db 841 CGCGAAGGATGATGACAAAGTCGATTGCGTCAACATCTTGTTCCGAAGTCCGGCGTC 900
 Qy 301 GGuAaIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
 Db 901 GAAGCTATCGTCGCGAAATGCGAGCCAAAGGCGAGAAAGTCCGCGCATCTGTTCCGC 960
 Qy 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
 Db 961 CACATCAAACTGAGAGCGGTCAACCTCGTAAATGTTTCGCGAGCAATTCGCGCAGAG 1020
 Qy 341 IleGlyAlaGluLysThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSer 360
 Db 1021 ATAGCGCGGAAAAAACCTCGGTACAAAATCGGGTACTTCGCGCGTCTGCTTC 1080
 Qy 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
 Db 1081 AACGTTGACGACATCGCTTGTATCAAAATCGTCGCGCATTCGCGGTGAGGCGCGTTC 1140
 Qy 381 ArgArgGluSerGlyValIleIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle 400

Db 1141 CCCCCGAGTCTGGCTGATCGGTCAAGCAAGACAGCGCAACCTGTTGCGTGGATC 1200
 Qy 401 GIupheProAgiIlelysglyglysPpopheAenIIeasPhrAptPpheaenSer 420
 Db 1201 GAGTTTCCCGCATCAAGGCGGCAACCGTTCAATATGACACCGACTGGTTCAATAGC 1260
 Qy 421 MetLeuSeGgiIleGlyGlnProLySGlyGlyLeValGluValSerHis 437
 Db 1261 ATGTTGAGGAATCGGCGACGCTTAAGCGGTAAGTGAAGTCAAGTCAAGCAC 1311
 RESULT 5
 AX398319 1311 bp DNA linear PAT 27-MAY-2002
 LOCUS Sequence 15 from Patent WO0220796.
 DEFINITION AX398319
 ACCESSION AX398319
 VERSION AX398319.1 GI:21261097
 KEYWORDS
 SOURCE MethyIomonas sp.
 ORGANISM MethyIomonas sp.
 Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
 Methylococcaceae; MethyIomonas.
 REFERENCE
 1 Koffas, M., Odum, J.M., Norton, K.C. and Ye, R.W.
 AUTHORORS Methanotrophic carbon metabolism pathway genes and enzymes
 TITLE Patent: WO 0220796-A 15 14-MAR-2002;
 JOURNAL E.I. DUPONT DE NEMOURS AND COMPANY (US)
 FEATURES
 source 1..1311
 Location/Qualifiers
 -/organism="MethyIomonas sp."
 -/mol_type="unassigned DNA"
 /db_xref="taxon:418"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,126-161 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-701-200-6 (1-437) x AX398319 (1-1311)
 Qy 1 AapValValThThrProGlyrHisLeuThrAlaAspIleAyrPheCySHSTPhePhe 20
 Db 1 GATGGTCAACATGCGCTTATCACTTAACGCGTGAATTCGATTGTCATGGTTTTT 60
 Qy 21 LeuAenPheAenPheTyThrLeuMetAenLySProLyLeValAlaIleLeuThra 40
 Db 61 CTTAACTTTAACTTCAACGCTCATGAACCAACCTTAAGGTTGCAATACAGACGA 120
 Qy 41 GlyGlyLeuAlaProCysLeuAenSerAlaIleGlySerLeuIleGluArgTyThrGlu 60
 Db 121 GCGCGGCTTGGCGCTTGTGAATTCGCAATCGGATGTTGATCGAACGTTATACGAA 180
 Qy 61 IleAenProSerIleGlyIleIleCysTyrArgGlyGlyTyThrGlyLeuLeuGly 80
 Db 181 ATGATCTTAAGCAATTAATCATTTGCTATCGCGGCGTTATTAAGCGCTGTGCTGGG 240
 Qy 81 AapSerTyProValThrAlaGluValArgLyLeValGlyValLeuGlnArgPheGly 100
 Db 241 GATTTCTTATCCAGTAAAGCGCGCAAGTGCCTTAAGGCGGCTTTCGCAAGCTTTGGG 300
 Qy 101 GlySerValIleGlyAenSerArgValLyLeuThrAenValLyAapCySValLyArg 120
 Db 301 GGTTCGTGATCGCAACAGCCCGTCAATTGACCAATGTCAAAGACTGCGTGAACGC 360
 Qy 121 GlyLeuValLySgIleGlyGlnAapProGlnLyLeValAlaAlaAapGlnLeuValLyAap 140
 Db 361 GGTTCGTGCAAAAGGATGAGATCGCAAAAGTGGCGGCTATCATTTGGTTAAGAT 420
 Qy 141 GlyValAapIleLeuHisThrIleGlyGlyAapAapThrAenThraIaAlaAapLeu 160

Db 421 GGTTGCAATATTCGACACCAATCGCGCGGATGATACCAATACGACAGCGGATTTG 480
 Qy 161 AlaAlaPheLeuAlaArgAenTyrglyLeuThrValIleGlyLeuProLySerThrVal 180
 Db 481 GAGCAATTCCTGGCCGAAATTAATTAAGCATGACCGCTCATTTGGTTTACTTAAGCCGT 540
 Qy 181 AapAenAapValPheProIleLySGlnSerLeuGlyValATrPThraIaIleGlnGly 200
 Db 541 GATTAACGATTAATTTCCATCAAGCAATCATAGTGTCTTGACCTCCCGCGAAGAGGC 600
 Qy 201 AlaArgTyrPheMetAenValAlaGluAenAenAlaAenProArgMetLeuIleVal 220
 Db 601 GCGCGTTATTTATCATGAACGTGGTGGCCGAAACCAAGCCACACCGCATGCTGATCGTA 660
 Qy 221 HisGlyValMetGlyValArgAenCysGlyTyThrPleuThraAlaThrAlaGlnGlyTyArg 240
 Db 661 CAGAAATGATGGCGCGTAACGTGGCTGCTGACCGCTGCAACCGCGAGAAATATCGC 720
 Qy 241 LyLeuLeuAapArgAlaGluTrPleuProGluLeuGlyLeuThrArgIleSerTyrgly 260
 Db 721 AATTTACTGACCGTCCGAGTGTTCGCGGAATGGTTGACTGTGATCTTATGAA 780
 Qy 261 ValHisAlaValPheValProGluMetAlaIleAapPleuGluAlaGlyAlyArgLeu 280
 Db 781 GTGACAGCGGATATTCGTTCCGGAATGGGATCGACCTGGAAGCCGAAAGCCGCTG 840
 Qy 281 ArgGluValMetAapLySValAapCySValAenIlePheValSerIleGlyAlaGlyVal 300
 Db 841 CCGAAATGATGACAAATGATTCATTCGTCACATCTTGTTCCAAAGTGGCGGCTC 900
 Qy 301 GluAlaIleValAlaGluMetGlnAlaLySGlyGlnGluValProArgAapAlaPheGly 320
 Db 901 GAAAGTATGTCGCGGAATATGACGCGCAAGCGCAAGAGTGGCGCATGCTTCGCGC 960
 Qy 321 HisIleLyLeuAapAlaValAenProGlyLyTrpPheGlyGlnPheAlaGlnMet 340
 Db 961 CACATCAAACTGATCGGTGCAACCTGGTAAATGTTGCGGCAAGCAATTCGCGCAGATG 1020
 Qy 341 IleGlyAlaGluLySerThrLeuValGlnLySeryGlyTyThrAlaArgAlaSerAlaSer 360
 Db 1021 ATAGGCGGCAAAACCTGATGACAAATAGGGAATGCTGCTGCTGCTTCC 1080
 Qy 361 AenValAapAapMetArgLeuIleLySeryCysAlaAapLeuAlaValGluCysAlaPhe 380
 Db 1081 AACGTTGACGACATGCTTGTGATCAATCTGTGCGCGCATTTGGCGGTCCAGTCCGCTTC 1140
 Qy 381 ArgArgGluSerGlyValIleGlyHisAapGluAapAenGlyAenValLeuArgAlaIle 400
 Db 1141 CCGCGGAGTCTGGCGTATCGGTCAAGCAAGACACGCGCAACGTTGCGTGGATC 1200
 Qy 401 GIupheProAgiIlelysglyglysPpopheAenIIeasPhrAptPpheaenSer 420
 Db 1201 GAGTTTCCCGCATCAAGGCGGCAACCGTTCAATATGACACCGACTGGTTCAATAGC 1260
 Qy 421 MetLeuSeGgiIleGlyGlnProLySGlyGlyLeValGluValSerHis 437
 Db 1261 ATGTTGAGGAATCGGCGACGCTTAAGCGGTAAGTGAAGTCAAGTCAAGCAC 1311
 RESULT 6
 AX398428 1311 bp DNA linear PAT 27-MAY-2002
 LOCUS Sequence 5 from Patent WO0220796.
 DEFINITION AX398428
 ACCESSION AX398428
 VERSION AX398428.1 GI:21261167
 KEYWORDS
 SOURCE MethyIomonas sp.
 ORGANISM MethyIomonas sp.
 Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
 Methylococcaceae; MethyIomonas.
 REFERENCE
 1 Koffas, M., Odum, J.M. and Schenle, A.
 AUTHORORS High growth methanotrophic bacterial strain MethyIomonas 16a
 TITLE

JOURNAL

Patent: NO 0220728-A 5 14-MAR-2002;
E.I. DUPONT DE MEMPHIS AND COMPANY, Legal Patent Records Center
(US)

FEATURES

Location/Qualifiers
1. 1311
/organism="Methylobacter sp."
/mol_type="unassigned DNA"
/db_xref="taxon:418"

ORIGIN

Alignment Scores:

Pred. No.:	2,126-161	Length:	1311
Score:	2284.00	Matches:	437
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-701-200-6 (1-437) x AX398428 (1-1311)

```

QY 1 AapValValThrTyrProTyrHisLeuThrAlaAspIleArgPheCysHisTyrPhePhe 20
DB 1 GATGGTGCACATGCGCCCTATCATTAACGGCTGATATTCGATTGTCATGCTTTT 60
QY 21 LeuAapPheAapPheTyrThrLeuMetAenLysProLysLysValAlaIleLeuThrAla 40
DB 61 CTTAACTTAACTTCTACACGCTCATGAACAACTTAAAGTTGCAATGCTGACAGCA 120
QY 41 GtGtGtLeuAlaProCysLeuAasSerAlaIleGtSerLeuIleGtUarGtYrThgU 60
DB 121 GCGGCGCTGGCGCCCTGTTTGAATTCGCAATCGGTAAGTTGATCGAAGCTTATACCA 180
QY 61 IleAapProSerIleGtUleIleCysTyrArgIleGtYrLysGtLysLeuLeuGt 80
DB 181 ATCGATCCAGCATGAATCATTTGCTATCGCGCGGTTATAAGCCCTGTTGCTGGGC 240
QY 81 AapSerTyrProValThrAlaGtUValArgLysLysAlaGtValLeuGlnAapPheGt 100
DB 241 GATTCTTATTCAGTAACGGCCGCAAGGCTTAAAGGCGGCTGTCGACGTTTGGC 300
QY 101 GtSerValIleGtLysAasSerArgValLysLeuThrAasValLysAapCysValLysArg 120
DB 301 GGTTCGTGATCGGCAACGCGCGCTCAATGACCAATGTCAAAGACTGCGTGAAGCC 360
QY 121 GtLysLeuValLysGtGtLysUapProGtLysValAlaAlaAapGtLysValLysAap 140
DB 361 GGTTCGTCAAGAGGCTGTAAGATCGCAAAAGTCCGCGCTGATCAATGCTTAAAGAT 420
QY 141 GtValAapIleLeuHisThrIleGtLysAapAapThrAasThrAlaAlaAapLeu 160
DB 421 GGTTCGATATTCGACACCATCGCGCGGATGATACCAATACGCGACGACGGAATTTG 480
QY 161 AlaAlaPheLeuAlaArgAasAasTyrGtLysLeuThrValIleGtLysProLysThrVal 180
DB 481 CGAGCATTCCTGGCCGAAATATTAACGACGACGCTATGTTTACTTAAACCGTC 540
QY 181 AapAasAapValPheProIleLysGtLysLeuGtLysValATrThrAlaAlaGtLysGt 200
DB 541 GATAACGACGATATTCGCGATCAAGCAATCATAGGTGCTTGACGCGCCGACGACGAC 600
QY 201 AlaArgTyrPheMetAasValAlaAlaGtUaAasAaAlaAasProArgMetLeuIleVal 220
DB 601 GCGCGTATTTATGATAACGTGGTGGCCGAAACCAACGCGCAACGCGCATGCTGATCGTA 660
QY 221 HisGtUaIleMetGtLysAasCysGtLysTyrLeuThrAlaAlaIleThrAlaGtLysTyrArg 240
DB 661 CACGAGAGTATGAGCGCGTAACTGCGGCTGCTGACCGCTGCAACCGCGAGAAATATCCG 720
QY 241 LysLeuLeuAapArgAlaGtLysTyrLeuProGtLysLeuGtLysLeuThrArgLysSerTyrGt 260
DB 721 AAATTAACGACGCGGATGCTGGTGGCGGAAATGGGTTGACTCGTGAATCTTATGAA 780
QY 261 ValHisAlaValPheValProGtUaMetAlaIleAapLeuGtUaIleAlaLysArgLeu 280

```

```

DB 781 GTGACGCGGATTCGTTCCGGAATGCGATCGACTCGAAGCCGAAGCCGACGCTTG 840
QY 281 ArgGtUaIleMetAapLysValAapCysValAlaIlePheValSerGtUaIleAlaGtLysVal 300
DB 841 CCGCAAGTATGACCAAGATCGATGCTGCAACATCTTCGTTCCGAAGTCCGCGCGTC 900
QY 301 GtUaIleIleValAlaGtUaMetGtLysAlaGtLysGtLysValProArgAapAlaPheGt 320
DB 901 GAAGCTATCGTCGGGAATGACAGCCCAAGGCGCAAGTGCCTGGCGATGCTTCCGCGC 960
QY 321 HisIleLysLeuAapAlaValAasProGtLysTyrPheGtLysGtLysPheAlaGtLysMet 340
DB 961 CACATCAACCTGATGCGGCTCAACCTGCTAATATGTTCCGCGACCAATTCGCGAGATG 1020
QY 341 IleGtLysAlaGtUaTyrThrLeuValGtLysSerGtLysTyrPheAlaArgAlaSerLys 360
DB 1021 ATAGCGCGGAAACCCCTGCTACAAATCGGGATGCTTCGCGGCTGCTTCCTTC 1080
QY 361 AasValAapAapMetArgLeuIleLysSerCysAlaAapLeuAlaValGtUaValAapPhe 380
DB 1081 AACGTTGACGACATGCGTTTATCATCAATCGCCCGCATTTGCGGTGAGTGCAGCTTC 1140
QY 381 ArgArgGtUaSerGtLysValIleGtLysAapGtUaAapAasGtLysValLeuArgAlaIle 400
DB 1141 CCGCGCGAGCTCGGCGTATGCTGTCACACGACGAAACGCGACGCTGTTGCTGATC 1200
QY 401 GluPheProArgIleLysGtLysLysProPheAasIleAapThrAapTyrPheAasSer 420
DB 1201 GAGTTCCGCCCATCAAGCGCGGCAACGTTCAATATCGACACGCACTGTTCAATAGC 1260
QY 421 MetLeuSerGtUaIleGtLysGtLysProLysGtLysValGtUaIleSerHis 437
DB 1261 ATGTTGACGGAATCGGCGACGCTTAAAGCGTAAAGTCAAGTCAAGCCAC 1311

```

RESULT 7

LOCUS BX294151/C

DEFINITION Pirellula sp. strain 1 complete genome; segment 19/24.

ACCESSION BX294151 BX119912

VERSION BX294151.1 GI:32446812

KEYWORDS complete genome.

SOURCE Rhodopirellula baltica SH 1

ORGANISM Rhodopirellula baltica SH 1

REFERENCE Gloeckner, F.O., Kube, M., Bauer, M., Teeling, H., Lombardot, T.,

Bacteria; Planctomycetes; Rhodopirellula.

1 (bases 1 to 295650)

Ludwig, W., Gade, D., Beck, A., Borzym, K., Heilmann, K., Rabus, R.,

Schlesner, H., Amann, R. and Reinhardt, R. Complete genome sequence of the marine planctomycete Pirellula sp.

strain 1

Proc. Natl. Acad. Sci. U.S.A. 100 (14), 8298-8303 (2003)

JOURNAL MEDLINE

PUBMED 22735913

12835416

2 (bases 1 to 295650)

Kube, M., Borzym, K., Heilmann, K., Klages, S., Marguardt, I.,

Lehrack, S., Beck, A., Pawlik, R., Reinhardt, R., Gloeckner, F.O.,

Bauer, M., Teeling, H., Lombardot, T., Ludwig, W., Gade, D., Rabus, R.,

Schlesner, H. and Amann, R. Direct Submission

Submitted (21-JAN-2003) Max Planck Institute for Molecular

Genetics, Proscience Innestrasse 73, D-4195 Berlin, Germany Max

Planck Institute for Marine Microbiology Celsiusstrasse 1, D-28359

Bremen, Germany

This project was carried out by

*Max Planck Institute for Molecular Genetics, Berlin, Germany; *Max

Planck Institute for Marine Microbiology, Bremen, Germany; in the

framework of the REXG-Project. http://www.regx.de -----

Center: Max Planck Institute for Molecular Genetics

Center code: MPING

----- Summary Statistics

Sequencing vector: pUC19; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 714284 bases at least Q40
Consensus quality: 714518 bases at least Q30
Consensus quality: 714548 bases at least Q20
Quality coverage: 8.03

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid sequence; assembly was additionally confirmed by long
range PCR and cosmid end sequences.

See <http://www.micro-genomes.mpg.de/pirella/> for more information
including minimal cloning path from a set of 220 cosmids out of
908. See the misc_feature tag below for the boundaries of the MTP
cosmid. ----- Annotation

Center: Max Planck Institute for Marine Microbiology
Celastrolase 1, D-26359 Bremen, Germany.

Center Code: MPIMM
Email: togempi-bremen.de
Phone: +49 (0)421 2028 938 Fax: +49 (0)421 2028 580

Three different programs (Glimmer, Critica and Orpheus) were used
for ORF-prediction. A nonredundant list of ORFs was generated by
mutable pairing of the results.
Automated annotation was done with the software package Pedant Pro
(<http://www.biomax.de>). All ORF predictions and annotations were
manually corrected by considering all results of the different
tools applied. See <http://www.regx.de> for more information and
access to supplementary information.

Location/Qualifiers

1. 295650

/organism="Rhodospirillum rubrum SH 1"

/mol_type="genomic DNA"

/strain="1"

/db_xref="taxon:243090"

/complement(414..1508)

/gene="pknH"

/locus_tag="RB10133"

/complement(414..1508)

/gene="pknH"

/locus_tag="RB10133"

/EC_number="2.7.1.1"

/function="unspecified signal transduction; unspecified
kinase or ATP dependent regulatory protein; cellular
communication/signal transduction"

/note="best DB hits: BLAST: swissprot:Q11053; PKH MYCN
PROBABLE SERINETHREONINE-PROTEIN; E=3e-35; embi: CAB94054.1;
(ALJ38672) putative serinethreonine-protein; E=1e-34
emb1:CA10713.1; (AJ132604) hypothetical protein
laccococcus; E=2e-33 COG: RV1266c; COG0515

Serine/threonine protein kinases: E=3e-36 PFAM: PF00069;
Protein kinase domain; E=7.6e-54"

/codon_start=1
/transl_table=1
/product="probable serine/threonine-protein kinase pknH"

/protein_id="CAD78719.1"
/db_xref="GI:32446813"

/translation="MMPHRLDFELGAVTGVCTGVYDGKILDDVHVAALNR
ODLAKKHPAVSODDLIARFRREVTIERLOHPNITIFPGSGDGLFTVMEKVD
GGTILDLLETNGLAMPVVDVARQVCSALQCAHNGVTRHDLKRGNTLELTDAVKL
GPGFARODSHDLTSGQITVGHVYMAPEQITGDEALSGKADLVAGCVLPEMLNR
VIFAGENFQALPEQHLRTKATPASIIVPVPELQVADCLKSGPDDRPVARSVOG
WVIEGKYNLAPDPSPVGNHSEGNAPKDVASDVTEKGRLLLEQIHVRLGSTS
REPVSAGKSLIVVAVLSGSGS"

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

CDS

/locus_tag="RB10134"
/function="ribosomal proteins"
/note="best DB hits: BLAST: gb:AAK05177.1; AE006340.10
(AE006340) 50S ribosomal protein L21; E=4e-10 pit:PF2920;
ribosomal protein L21 U212 (imported) - Ureaplasma;
E=3e-11 gb:AAK35614.1 (AF041468) ribosomal protein L21
(guilardi); E=3e-10 COG: RV2442c; COG0261 Ribosomal
protein L21; E=3e-07 PFAM: PF00829; Ribosomal prokaryotic
L21 prot; E=2.2e-36"
/codon_start=1
/transl_table=1
/product="50S ribosomal protein L21"
/protein_id="CAD78720.1"
/db_xref="GI:32446814"
/translation="MVAFVDSGRQRYVEPGMELDVYDIAAGENLKFEYLVAGAE
DGLKIGAPTLDSVSASVLSGSDKITYIQFRRRKRKSKRTGHRKXNTRIRIEIA
GV"
/complement(1767..2018)
/locus_tag="RB10135"
/complement(1767..2018)
/locus_tag="RB10135"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAD78721.1"
/db_xref="GI:32446815"
/translation="MLDRPIFGSTDLFTSDLLSLAKNOPMRSKRVSVILLDSKIR
PSCGPPLHPEPSRHVCHCRMTWTPSPARNGTRRLP"
2019..2177
/locus_tag="RB10136"
2019..2177
/locus_tag="RB10136"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAD78722.1"
/db_xref="GI:32446816"
/translation="MLVLRNRSLSVGSAGAAARCVTGIAPNGTYFTPGGCEP
IRASPVTP"
/complement(2151..2372)
/locus_tag="RB10137"
/complement(2151..2372)
/locus_tag="RB10137"
/codon_start=1
/transl_table=1
/product="hypothetical protein-putative conserved
hypothetical protein"
/protein_id="CAD78723.1"
/db_xref="GI:32446817"
/translation="MSRSRFSGVLYDAPLPKRSLGGRATPFRTRGRGSPMENGADC
PPPEISLNRPTPPAPARAGFSRRWRCT"
/complement(2478..3281)
/locus_tag="RB10138"
/complement(2478..3281)
/locus_tag="RB10138"
/function="inner membrane organization"
/note="best DB hits: BLAST: pir:134929; hypothetical
protein SC3P9.07 SC3P9.07 - Streptomyces; E=0.018"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein-putative membrane
lipoprotein"
/protein_id="CAD78724.1"
/db_xref="GI:32446818"
/translation="KRLVDFPNAFPFLQRNSTSSALNSLVFAGCCILAVGFEISTAE
DEPAQTSVNSYKIFDTNAPGNRQDGDVFNVDGDDTLTWEGTALSGSPITGTR
TNFVYNFELVPTMGHLKPKAGNSGVFAWPMKALDLPNNRLPNTGIEVQMDLDVGR
NTEYTKGPELPTMTSGDIFAVGKSMQPPPLSPDGRHSFSAETTNPGEWNYVY
RGINEIRLRMNWGVSVSGRSCSPDGLFCLSEDSGPIRFREIMRELPE"
/complement(3382..3621)
/locus_tag="RB10139"
/complement(3382..3621)
/locus_tag="RB10139"

```

/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAD78725.1"
/db_xref="GI:32446819"
/translation="MRHQRDRKTRIGISKNPILPSRYRGRVLRKRRSRPSFLRE
SRITLISHGSRIRHLAPARGRCHNPSSTHQKS"
3707..5527
/locus_tag="RB10140"
3707..5527
/function="anion transporters (Cl(-), SO4(2-), PO4(3-),
etc.)"
/notes="best DB hits: BLAST: pir:B82127, sulfate permease
family protein VC2031 [imported] -; B=1e-89 pir:G83325;
probable sulfate transporter PA2563 [imported] -; B=5e-53
dbj:BA804133.1; (AP001508) sulfate permease [Bacillus;
E=1e-51 COG: VC2031, COG0659 sulfate permease and related

```

Alignment Scores:

```

Pred. No.: 2,286-102 Length: 295650
Score: 1531.50 Matches: 286
Percent Similarity: 79.95% Conservative: 61
Best Local Similarity: 65.90% Mismatches: 78
Query Match: 67.05% Indels: 9
DB: 1 Gaps: 1

```

US-10-701-200-6 (1-437) x BX294151 (1-295650)

```

OY 4 ThrTPRProTYHLeuThraAlaAriLeaRphCyHLeaTRPheRheLeuAmphe 23
DB 281440 TCTTGAAATTCACACGCGTCCGCTTGTAGCTTCGAAACCTATCTCTCAATC 281381
OY 24 AenPheTYrThrLeuMeCAmLyAerProLyAAlaLeuThraAglYleu 43
DB 281380 -----AAAGCTGTGGAATTCACGCGTGTGGCTTG 281348
OY 44 AlaProCYLeuAenSerAlaIleegLYserLeuIleGUAATGYrThrGluIleAmpPro 63
DB 281347 GCCCGGTGTCTGTCACTCGGCCATCGAGCTTGATGTAAGCTTACACAGAGCAAGCTCT 281288
OY 64 SerIleGluIleIleCYeTYrAglYleuTYrLYeGlyLeuLeuLeuGlyAmpSerTYr 83
DB 281287 GAATTCGAATCTCTGCTCAACGCTGTGTACAAAGCTTCTCTCGGCGACAGCTTC 281228
OY 84 ProValThraAgluValAglYleuAglYleuGlnAglPheGlyYleuSerVal 103
DB 281227 GTCGTGATGCCCATTCGCGGAAGACGGCGCATCTTGCCACAGATGTGTGCGACCCC 281168
OY 104 IleGlyAenSerAryValIleuThraAenValIlyAerCYeValIlyAerGlyLeuVal 123
DB 281167 ATCGGCAACAGCCGGGTCAAACTGACCAACGTTGCGATGTGTGAAACGTGTGGTTC 281108
OY 124 LybGluIleGlyAerProGlnIlySerValAlaAlaAerGlnLeuValIlyAerGlyValAer 143
DB 281107 AGCGAAGACAGATCTCACTGACAGTGTGCGCGAGCGATTCCATTCGACAGAACTGCAC 281048
OY 144 IleLeuH1eThrIleGlyYleuAerPThraAenThraAlaAlaAerPheuAlaAerPhe 163
DB 281047 GGTGTGACACACCATCGGTGTGACACACCAACACCGGCGGAGCTTGCGCGCTTAT 280988
OY 164 LeuAlaAerAenAerTYrGlyLeuThraValIleGlyLeuProLYeThraValAerAenAer 183
DB 280987 TTGGCCAGACAGCAATATCACTGACGAGTGTGTGTTGCCCAAGCATTCACACAGAC 280928
OY 184 ValPheProIleLYeGlnSerLeuGlyAlaTPThraAlaAglGlnGlyValAerTYr 203
DB 280927 GTCATCCCATCAACAAAGCTTTGGTGTGTGACCGCGCGGAGAGAGCCCAAGTTC 280868
OY 204 PheMeCAmValValAgluAenAenAaAerProAryMeLeuIleValH1eGlyVal 223
DB 280867 TTGAAATATGTGTGTGCGGACACACACGCAATCCAGCATCTGATTCATTCAGAGTTC 280808

```

```

OY 224 MetGlyAerAenCYeGlyTYrPLeuThraAlaAlaThraAglGlnGlyTYrAryLYeLeu 243
DB 280807 ATGGGTCCCAACTGTGTGTGTGTGACCGCCCGCACCGCAAGCAATATCGCAAGCTTGG 280748
OY 244 AenAryAgluTYrPLeuProGlyLeuGlyLeuThraAryGlySerTYrGlyValH1aAla 263
DB 280747 CAGAAAGCTCAACTTCTTCCCTGAAATGGGACTCACTCAAGAACGTCGTGACCTGCACGCG 280688
OY 264 ValPheValProGluMeAlaIleAerPLeuGluAgluAlaLYeAryGlyLeuAryGlyVal 283
DB 280687 GTCTTTGCTTCGATGATGATTTTGACCTCGAAGAAAGCAACGTTTGGCGCGCATC 280628
OY 284 MetAerLYeValAerCYeValAenIleAerValSerGlyValAglYValGluAlaIle 303
DB 280627 ATGATGTGATCGATGTGTGCTCAACATCTTCCTCGAAGGTGTGTGCGCTGCACACCATC 280568
OY 304 ValAgluMeAlaIleLYeGlyYleuGlnGlyValProAryAerAlaPheGlyH1eIleLYe 323
DB 280567 GTCAAGAAATGGAATCAACGCGCGCAAGCGTCCCAAGATGCATTCCGCGCATACAA 280508
OY 324 LeuAerAlaValAenProGlyLYeTPRheGlyGlnGlnPheAlaGlnMeIleGlyAla 343
DB 280507 CTGACGACAGTGAACCCAGCAAAATGTTCGCAAAAGTTCGTGACATGTGTGGGTGCC 280448
OY 344 GluLYeThraValGlnLYeSerGlyTYrPheAlaAryAlaSerAlaSerValAer 363
DB 280447 GAAAAACGCTTGTCAAAGAGCGGTACTTCAGTCTGTGCGCGCGCCCAACCGGAA 280388
OY 364 AenMeCArGlyLeuIleYerSerCYeAlaAerPheuAlaValGluCYeAlaAerAryAglu 383
DB 280387 GACATCGCGCTGATCGGTGTGTGTGTCGAAGAGCGTCGAATGTGCAATCGAGGCGCATC 280328
OY 384 SerGlyValIleGlyH1eAerGlyAerPheAenGlyAenValIleuAryAlaIleGluPhePro 403
DB 280327 GGTGTGTGTGTGCGGAAGACAGATCAAAACAGAGCTTCGTGCGCATTCGATTTGAA 280268
OY 404 ArgIleLYeGlyLYeAerProPheAenIleAerThraAerTPRheAenSerMeIleuSer 423
DB 280267 CGCATCGCGCGGCGGAAACCTTTCACATCAACGTAAGATGTGTGTGACCTGCTTCA 280208
OY 424 GluIleGlyGlnProLYeGlyLYeValGluValSerH1e 437
DB 280207 CAGATGGCGCAACCAAGCGCAAGTCTGGAAGCTGTCA 280166

```

RESULT 8
AF246209 1427 bp mRNA linear INV 21-AUG-2003
LOCUS Maestigamoeba balamuthi Pp1-phosphofructokinase (Pp1pfk) mRNA,
DEFINITION complete cds.
ACCESSION AF246209
VERSION AF246209.1 GI:7862072
KEYWORDS
SOURCE Maestigamoeba balamuthi
ORGANISM Maestigamoeba balamuthi
Eukaryota; Pelobiontida; Maestigamoebidae; Maestigamoeba.
REFERENCE 1 (bases 1 to 1427)
AUTHORS Muller,M., Lee,J.A., Gordon,P., Gaasterland,T. and Sensen,C.W.
TITLE Presence of prokaryotic and eukaryotic species in all subgroups of
the Pp(1)-dependent group II phosphofructokinase protein family
J. Bacteriol. 183 (22), 6714-6716 (2001)
JOURNAL 21528850
MEDLINE 11673446
PUBMED 2 (bases 1 to 1427)
REFERENCE Muller,M., Lee,J.A. and Sensen,C.W.
AUTHORS Direct Submission
TITLE Submitted (15-MAR-2000) Laboratory of Biochemical Parasitology, The
JOURNAL Rockefeller University, 1230 York Avenue, New York, NY 10021, USA
FEATURES
source
1..1427
/organism="Maestigamoeba balamuthi"
/mol_type="mRNA"
/db_xref="taxon:108607"
1..1427
gene


```

DB: 6 Gaps: 1
US-10-701-200-6 (1-437) x CQ363740 (1-8530)
QY 33 LysLysValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAmsSerAlaIleGly 52
DB 3962 AAGAAAGTCGACCTCTCCACCGGGCGGATTTGGCCCTGCTTATCCACCGCATCACT 4021
QY 53 SerLeuIleGluArgTyrThrGluIleApsProSerIleGluIleIleCysTyrArgGly 72
DB 4022 GGATTGATCCAGCGCTACACGAGGAGACCTGAGTGCATGATCATGCTTACCAAGC 4081
QY 73 GlyTyrGlyLeuLeuLeuGlyApsSerTyrProValThrAlaGluValArgLys 92
DB 4082 GGCTACGAGGGCTCTCTCAAGGAGACTTCTCGAGTCCAGCACCGCTTGCAGAGAT 4141
QY 93 AlaGlyValLeuGluApsPheGlyGlySerValIleGlyAmsSerArgValLysLeuThr 112
DB 4142 GCCGAGATCTCTAAGCGTTTGGTGGATCCCGATTGGCACTCGCGGCTCAAGCTGACC 4201
QY 113 AsnValLysApsCysValLysArgGlyLeuValLysGlyGluApsProGluLysVal 132
DB 4202 AAGCGCGCGACCTGTAAAGCGTGTGTGTGCGCGAAGCGAGACCGCTCAAGGTT 4261
QY 133 AlaAlaApsGluLeuValLysApsGlyValApsIleLeuHisThrIleGlyLysAps 152
DB 4262 GCCGCGCATCGCTGTGTGCGCGAGGCTTGCATTTGCACACATCGGTGTGACGAC 4321
QY 153 ThrAenThrAlaAlaApsLeuAlaApsLeuAlaApsLeuAlaApsAmsTyrGlyLeuThr 172
DB 4322 ACTAATACCAACCGCGCCGATCTGGCCGCTTATCTGCTGAGAAATCACTACGCTCAC 4381
QY 173 ValIleGlyLeuProLysThrValApsApsApsValPheProIleLysGlnSerLeuGly 192
DB 4382 GTCGTGGCTGCGCGAAGCATCATGTAACGATGTCGATTCGATTCGATTCGATTCGAT 4441
QY 193 AlaTyrThrAlaAlaGlyGlnGlyAlaArgLysPheMetApsValValAlaGluAms 212
DB 4442 GCTTGAGCGGCTGCGAAGCATCATGTAACGATGTCGATTCGATTCGATTCGATTCGAT 4501
QY 213 AlaApsProApsMetLeuIleValHisGluValMetApsValApsApsGlyTyrLeuThr 232
DB 4502 TCCGATACCGATGCTCATCTGCTCCAGAGTCAAGGCTGAGCTGCGCTCAC 4561
QY 233 AlaAlaThrAlaGlnGlyTyrArgLysLeuApsApsArgAlaGluTyrLeuProGluLeu 252
DB 4562 GCCGCTACCGCGCAATATACCGCAATGCTCGATTCACACAGTGGCTACCTGAGATC 4621
QY 253 GlyLeuThrArgLysSerTyrGluValHisAlaValPheValProGluMetAlaIleAps 272
DB 4622 GGGCTGTCCAGAGGCTGCGATGCTCCAGCGCTATATGTCCTCCAGGCTCATCATCGAC 4681
QY 273 LeuGluAlaGluAlaLysArgLeuArgGluValMetApsLysValApsPysValAmsIle 292
DB 4682 CTTGAGCGGAGCGCGACGCTCTCAACAGGTCATGAGAGAGTTGGCAAGCTCACTATC 4741
QY 293 PheValSerGlyGlyAlaGlyValGluAlaIleValAlaGluMetGlnAlaLysGlyGln 312
DB 4742 TTCCTGTCCAGAGGCTGCGATGCTCCAGCGCTATATGAGAGATGGAAGAGCGCTGAG 4801
QY 313 GluValProApsApsAlaPheGlyHisIleLysLeuApsAlaValApsProGlyLysTyr 332
DB 4802 GAAAGTCCCGCGGATCTCTTGTGCTCAAGCTCGACAGCTGAGAACCCGGGTGCTTGG 4861
QY 333 PheGlyGluGlnPheAlaGlnMetIleGlyValGluLysThrLeuValGlnLysSerGly 352
DB 4862 TTCGCGAAGCAAGTTCCCGCAAGCTCGGTTGCCGAAAGGTCAATGCTCCAGAGTCCGAC 4921
QY 353 TyrPheAlaArgAlaSerAlaSerAmsValApsApsMetArgLeuIleLysSerCysAla 372
DB 4922 TACTTCTCCGCTTCCGCAAGCTCGAAGAGCTGAGCTGATCATTCGCGCGCTGACT 4981
QY 373 ApsLeuAlaValGluCysAlaApsApsArgGluSerGlyValIleGlyHisApsGluAps 392

```

```

DB 4982 GACCTCGCTGCTGATGCGCTTGGCCGCTGAGACTGCGCTTATGCGCCAGAGAGAG 5041
QY 393 AmsGluAmsValLeuArgAlaIleGluPheProArgIleLysGlyGlyLysProPheAms 412
DB 5042 AAGCGCGACAGTGTACAGAACATGCGCTTGCATGCGGATCAAGAGGCGCAAGCTTTCAC 5101
QY 413 IleApsThrApsTyrPheAmsSerMetLeuSerGluIleGlyGln----- 427
DB 5102 ACCATCAAGCGCTGCTTACCGCCATCTCAAGCATCGGCGAGCTGAGCTATGG 5161
QY 428 ProLysGlyGlyLysValGlu 434
DB 5162 CCAACAGTGGCTTCCGAA 5182

RESULT 10
AE017283_11/c
WPCOMMENT
Sequence split into 26 fragments LOCUS AE017283 Accession AE017283
Fragment Name Begin End
AE017283_01 1 110000
AE017283_02 100001 210000
AE017283_03 200001 310000
AE017283_04 300001 410000
AE017283_05 400001 510000
AE017283_06 500001 610000
AE017283_07 600001 710000
AE017283_08 700001 810000
AE017283_09 800001 910000
AE017283_10 900001 1010000
AE017283_11 1000001 1110000
AE017283_12 1100001 1210000
AE017283_13 1200001 1310000
AE017283_14 1300001 1410000
AE017283_15 1400001 1510000
AE017283_16 1500001 1610000
AE017283_17 1600001 1710000
AE017283_18 1700001 1810000
AE017283_19 1800001 1910000
AE017283_20 1900001 2010000
AE017283_21 2000001 2110000
AE017283_22 2100001 2210000
AE017283_23 2200001 2310000
AE017283_24 2300001 2410000
AE017283_25 2400001 2510000
AE017283_26 2500001 2600000

Continuation (12 of 26) of AE017283 from base 1100001 (AE017283 Proplionibacterium acnes)

Alignment Scores:
Pred. No.: 1,21e-95 Length: 110000
Score: 1434.50 Matches: 273
Percent Similarity: 80.34% Conservative: 54
Best Local Similarity: 67.08% Mismatches: 75
Query Match: 62.81% Indels: 5
DB: 1 Gaps: 1

US-10-701-200-6 (1-437) x AE017283_11 (1-110000)
QY 33 LysLysValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAmsSerAlaIleGly 52
DB 3962 AAGAAAGTCGACCTCTCCACCGGGCGGATTTGGCCCTGCTTATCCACCGCATCACT 84119
QY 53 SerLeuIleGluArgTyrThrGluIleApsProSerIleGluIleIleCysTyrArgGly 72
DB 4022 GGATTGATCCAGCGCTACACGAGGAGACCTGAGTGCATGATCATGCTTACCAAGC 84059
QY 73 GlyTyrGlyLeuLeuLeuGlyApsSerTyrProValThrAlaGluValArgLys 92
DB 4082 GGCTACGAGGGCTCTCTCAAGGAGACTTCTCGAGTCCAGCACCGCTTGCAGAGAT 83999
QY 93 AlaGlyValLeuGluApsPheGlyGlySerValIleGlyAmsSerArgValLysLeuThr 112
DB 4142 GCCGAGATCTCTAAGCGTTTGGTGGATCCCGATTGGCACTCGCGGCTCAAGCTGACC 83939

```

```

Qy 113 AenVallyAspCyVallyAspArgglyLeuVallyGluGlyGluAspProGluVal 132
Db 83938 AACGCCGCGACCTCTAAAGCGTGTGTCGCGCAAGGCGACGCGCTCAAGGTT 83879
Qy 133 AlaAlaAspGluLeuVallyAspArgglyValAspIleLeuHisThrIleGlyGlyAspAsp 152
Db 83878 GCGCGCGATCGTGTGTTCCGACGCGGGGTGACATTTTCACACCATTCGTGTGACGAC 83819
Qy 153 ThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnThrGlyLeuThr 172
Db 83818 ACTAATACCAACCGCCCGCATCTGCGCGCTATCTGGCTGAGAAATACACGCGCTCAC 83759
Qy 173 ValIleGlyLeuProlyThrValAspAsnAspValPheProIleGlyGlnSerLeuGly 192
Db 83758 GTCGTTGGCTGCGCGCAAGACATCGAATACGACGTGTCGCGCATTCGTCATCGTAGGT 83699
Qy 193 AlaThrThrAlaAlaGluGlnGlyAlaArgTyPheMetAsnValAlaGluAsnAsn 212
Db 83698 GCTTGGACGCGCTGCGCAAGGTTCCCGCTTGTTCAGAAACATCGTTGGCGACCAAC 83639
Qy 213 AlaAsnProArgMetLeuIleValHisGluValMetGlyArgAsnCyAspGlyThrLeuThr 232
Db 83638 TCCGATTCACGCGATGCTCATCTGTCACGAGTCAAGCGCTGACTGCGCTGCTCAC 83579
Qy 233 AlaAlaThrAlaGlnGlyTyArgTyLeuLeuAspArgAlaGluThrLeuProGluLeu 252
Db 83578 GCGGCTACCGCGCCCAATACCGCGAATGGCTGATACACAGTGGCTGACTGAGATC 83519
Qy 253 GlyLeuThrArgGluSerTyGluValHisAlaValPheValProGluMetAlaIleAsp 272
Db 83518 GGGCTGTCCAGAAAGCGCTGCGATGTCCACGCGGTATGTCTCCCGACGCTCACATCGAC 83459
Qy 273 LeuGluAlaGluAlaValAspArgGluValMetAspLeuValAspCyValAsnIle 292
Db 83458 CTTGAGCGCGAAGCCCGACGCTCAACAGGTATGACAGAGTGGCAACGCTACTATC 83399
Qy 293 PheValSerGluGlyValGlyValAlaValAlaValAlaGluMetGlnAlaGlyGln 312
Db 83398 TTCCTGTCCGAGGGGCTGGCGCTGACGCGCATATTGAGAGATGGAAGAAAGCGTCA 83339
Qy 313 GluValProArgAspAlaPheGlyHisIleValLeuAspAlaValAsnProGlyVal 332
Db 83338 GAAGTCCCGCGCATCTTGTGTCACTGTCAGCAAGCTGCAAGCTGCAAGCTGCTTGG 83279
Qy 333 PheGlyGluGlyPheAlaGlnMetIleGlyValGluValThrLeuValGlnValSerGly 352
Db 83278 TTCGGCAACAGTTCGCGCAAGCTGCGTGGCGCAAGAGTCTGTCGCAAGTCCGCG 83219
Qy 353 TyrPheAlaArgAlaSerAlaSerAsnValAspAspMetArgLeuIleValSerCyAla 372
Db 83218 TACTTCTCCGTTCCGACGCTCGAAGCAAGCTGAGAGTCACTCCGCGCTCACACT 83159
Qy 373 AspLeuAlaValGluCyAlaPheArgArgGluSerGlyValIleGlyHisAspGluAsp 392
Db 83158 GACCTGCTGTGATGCGCGCTGCGCGCGGTAAAGTACGCGCTTATCCGCGCAGCGAGAG 83099
Qy 393 AsnGlyAsnValLeuArgAlaIleGluPheProArgIleGlyGlyValAspProPheAsn 412
Db 83098 AACGGGCAACCGTTGCAAGCAATGCGCTTGATCGATCAAGGCGCGCAAGCTTTCGAC 83039
Qy 413 IleAspThrAspTrpPheAsnSerMetLeuSerGluIleGlyGln----- 427
Db 83038 ACCACTCAGCGCGGTTCACCGCGCATGCTCAGCGAATCGGCGAGCGCTGACAGCTATGG 82979
Qy 428 ProGlyGlyGlyValGlu 434
Db 82978 CCCAACAGTGGCCCTCCGAA 82958

```

```

RESULT 11
PRSPFK PRSPFK 1400 bp DNA linear BCT 24-MAY-1993
LOCUS P.freudenreichii pyrophosphate-fructose-6-phosphate
DEFINITION

```

```

ACCESSION M67447
VERSION M67447.1 GI:150930
KEYWORDS pyrophosphate-fructose 6-phosphate 1-phosphotransferase.
SOURCE Proionibacterium freudenreichii
ORGANISM Proionibacterium freudenreichii
REFERENCE 1 (bases 1 to 1400)
AUTHORS Lador, U.S., Gollapudi, L., Tripathi, R.L., Latshaw, S.P. and Kemp, R.G.
TITLE Cloning, sequencing, and expression of pyrophosphate-dependent phosphofructokinase from Proionibacterium freudenreichii
JOURNAL J. Biol. Chem. 266 (25), 16550-16555 (1991)
MEDLINE 91358443
PUBMED 1653240
COMMENT Original source text: P. freudenreichii DNA.
FEATURES
source
location/Qualifiers
1..1400
/organism="Proionibacterium freudenreichii"
/mol_type="genomic DNA"
/db_xref="taxon:1744"
74..1288
/gene="PFK"
/gene="PFK"
/EC_number="2.7.1.90"
/codon_start=1
/transl_table=1
/product="pyrophosphate-fructose 6-phosphate
1-phosphotransferase"
/protein_id="AAA25675.1"
/db_xref="GI:150931"
/translation="MWKVALITAGGAPACLSAATLAKRTYEVSPETLLIGRYGY
EGLGDSLESPRAHAYDR.FSFGSGSIGSRVKT.VNVKLVARGIVASGDDPKV
AADQIADGVADVLRITGGDTTTLADLAAYLRQHDYPLTVVGLPRTTINDYPIKOS
LGAWTRADEGARFANVIAENHNAAPRELIIHIMGNCGYLAETSRRYVAMVDAQOM
LEBAGIDRGMIDIHLYVEARTIDDAEERLRTYDVEGSAVNIPISEBAGVADIVAO
MGATGCEVPTDAFGHVDKINPGAMFAQFAPERIGAGTMYOKSGYFERSASNNAD
LELIATATMAVDALAGTPGVGQDEBAGDKLSVIDFRIGIAGHKRFIDTLDPYQLL
ARIGQAPAPFAAA"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 6,23e-89 Length: 1400
Score: 1314.00 Matches: 248
Percent Similarity: 77.78% Conservative: 60
Best Local Similarity: 62.63% Mismatches: 88
Query Match: 57.53% Indels: 0
DB: 1 Gaps: 0

```

us-10-701-200-6 (1-437) x PRSPFK (1-1400)

```

Qy 33 LybLybValAlaIleLeuThrAlaGlyGlyLeuAlaProCybLeuAsnSerAlaIleGly 52
Db 80 AAAAAGGTGCTCTGCTGACCGCGTGTGCTTGCCTCGTCTTCTTCCTGCGGCATCGCT 139
Qy 53 SerLeuIleGluArgTyThrGluIleAspProSerIleGluIleIleCyTyArgGly 72
Db 140 GAGCTCATCAAGGCTATACCGAGGTATCAACCGAAACGACCTCACTGCGCTATTCCTAT 199
Qy 73 GlyTyLeuGlyLeuLeuLeuGlyAspSerTyProValThrAlaGluValArgTybLyb 92
Db 200 GCGTATGAGGCGCTGCTCAAGGCGATTCCTCGAATTCCTCCCTGCGCGCGACAC 259
Qy 93 AlaGlyValLeuGlnArgPheGlyGlySerValIleGlyAsnSerArgVallybLeuThr 112
Db 260 TACGACCGGCTCTTACGTTCCGCGGGTCAACGATCGGGAACCTCCGCGGTCAAGCTCAC 319
Qy 113 AenVallyAspCyVallyAspArgglyLeuVallyGluGlyGluAspProGluVallyVal 132
Db 320 AATGTAAAGACTGTGTCGCGGGGCTGTTGCTTCCGCGCATGATCCCTCAAGGTT 379

```


/product="PUTATIVE TRANSCRIPTION REGULATOR PROTEIN"
/protein_id="CAC46509.1"
/db_xref="GI:15074952"
/db_xref="GOA:Q92P63"
/labelation="MTGMDKGTKNRATTAAPVEGGRLTKLTIAFTWGLGVTVSRALK
DADPDGAETKEVERLVRAIOIGYOPRRARLRITGETNVISVLTLLEEIMGITSPMVI
RTTILLAGOYHLYVTTPSSADPLGRYLTDGCANAGVISTEIBNDPRVILLTER
RIPTTHGTRTENGLIHPHYDFDNERRFAIEAVRKLDNRKRLLVLEPPNLIFFSHKK
TGFEVGVDPFGEAAVSFHOVNIDSLVAIRAFEXLMKPSPAIPGVVSGSQAIALII
AGIEAKGVGEDAEAWSKVPDPFLMRPEVMWTYEDIRIAGEELAKVJIEGRP
PETLOLSPOEFQQPMARSAYS"

repeat_region

Gene

CDS

complement(3036..3302)
/gene="SMC04261"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypotheical
predicted by Codon usage
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="HYPOTHEICAL PROTEIN"
/protein_id="CAC46510.1"
/db_xref="GI:15074953"
/db_xref="UniProt/TREMBL:Q92P62"
/labelation="MKFTVIEEGKAAYTAASVTIRDEARRIKAAGYEQWRRSILA
TSAPIPGPYOLXKOIDPVAGLEQLNPDISPDFDKKWPAAGI"
complement(3407..4837)
/gene="gnd OR SMC04262"
complement(3407..4837)
/gene="gnd OR SMC04262"
complement(3407..4837)
/EC_number="1.1.1.44"
/function="small molecule metabolism; energy metabolism,
carbon; oxidative branch; pentose pathway"
/note="Product confidence : probable
Gene name confidence : putative
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="PROBABLE 6-PHOSPHOGUCONATE DEHYDROGENASE
(DECARBOXYLATNG) PROTEIN"
/protein_id="CAC46511.1"
/db_xref="GI:15074954"
/db_xref="GOA:Q92P61"
/db_xref="UniProt/TREMBL:Q92P61"
/labelation="MSOAELICLGWGWSMLANIABEKGNRIAVFNRYDATCKFYNA
EGALKDQIVPCETEEFPVALAIRPRPIIMIKGDDPVDOOMALKPKIAGKLINDIPA
GANARDTRMPREDALKDSGLTFICMGVGSSEGBAHGPSIVGCTESSVRREVVELYS
IANKUDSPCARULEBNAGHFVKTIHNHGIEADMOKMAEITYGILRDCLKTADOTIGE
VFANNKRNLNSYLIETEEKVKADLPITGFRPYDMILIUKAQQGKTGWSTEAONONGK
VPATTIEAVNASISSAKEERBAEKYLG.PVGEIKVADRDSFIKDLNALALAARIKI
GVAYAGCFVVAAAASKREFGMNIMPITIAXIWAGCIIRSOFDELITTATAKPDAANIIL
VYPAPAANYKVSDGALRVYSTAVVIGLPVALLASALGYFDYSRGRCGTNAVIQAGDR
FGCAJGFPRVOADSHHGPWGSLANA"

Gene

CDS

5273..6691
/gene="SMC04263"
5273..6691
/function="cell processes; transport of small molecules;
amino acids, amines, peptides"
/note="Product confidence : putative
Gene name confidence : hypothetcal

predicted by Codon usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=1
 /product="PUTATIVE AMINO ACID CARRIER TRANSMEMBRANE
 PROTEIN"
 /protein_id="CAC46512.1"
 /db_xref="GI:15074955"
 /db_xref="GOA:Q92P60"
 /db_xref="UniProt/TREMBL:Q92P60"
 /translation="MDPTIGFLNTIFMCKVYLIGLVALVETVALGPPQIHSEMF
 RVLSRSGSKDAAGISPPFOALVSLASRGTGNLAVVALVIGGGATLPMHIVAFV
 MATVAVSSALALQLYKIRNEDQYRGSPFVYFARGINAAWALITFACILISGLFVNA
 VQNSIADAVOAGFVPLTAVAGVALVSLGVIPEGIIQIAVAIVVFPMAAYLLT
 AVYLLINADVAGAPVLTWTIISSAFGSCAGGITGGIAAAMGVKRGIFSEADGWS
 APNIAAVATPVPHHSSQGFQVSGFIDTILICATSMILLSGTLEPSSGVGTQOL
 TQAMSTHIGSAGTYFLTAIFPFPAFTSIIGNVSYAENALTYLGGGNRLGLIMRCAT
 LAMVWKAYESITTFPDNADASMGLMATTINLVAILLSGTIAKLTQDYFQRKAGAVP
 VFHADYPDLQKIDGELWSRD"
 6773.7663
 /gene="SMC04264"
 6773.7663
 /gene="SMC04264"
 /function="miscellaneous; not classified regulator"
 /note="Product confidence : putative
 Gene name confidence : hypothetical
 predicted by Codon usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=1
 /product="PUTATIVE TRANSCRIPTION REGULATOR PROTEIN"
 /protein_id="CAC46513.1"
 /db_xref="GI:15074956"
 /db_xref="GOA:Q92P59"
 /db_xref="UniProt/TREMBL:Q92P59"
 /translation="MGTSQPAVYDADRAPBELQOEARLGLPALAQTSGIYRQSP
 AALAFPGKWLHMPDGPASIVAVPDCSDIIVSSKGLAIVGDRILAAPELPAG
 TIIQAPRIGAAAMIRPLAEMTQGVPLHLWGPANETEARIRDAASPERIGVL
 AALLGRLSAAPPPTIACVAHLQVRSITVDADIRTLVREIGISERTLRNHCHELT

Oy	133	Aa AaAaBpGlnLeuValIleAaBpGlyValAaP IleuHisThrIleGlyValAaAaP	152
Db	273860	GGCGCGGAAGGCTGCGCTCCACAGGACATCTCTCACACGATCGCGCGACGAT	273919
Oy	153	ThrAenThrAlaAlaAaBpLeuAlaAaPheLeuAlaAaBpAaBnThrGlyLeuThr	172
Db	273920	ACCAACACACCGCGCGCATCTCCGAGCTATCTCCGGGGCCAAAGCGCTACATCGAAG	273979
Oy	173	ValIleGlyLeuProIleThrValAaBpAaBpValPheProIleIleGlnInsertIleGly	192
Db	273980	GTCGCGGCTGCTCCCAAGACGGTCTGCACACGATGTGCTCCATCCGCGACGCTCGGC	274039
Oy	193	AlaTPThrAlaAlaGlnGlnIleAaGlyrPrpMetAaBnValAlaGluAaBnAa	212
Db	274040	GCTTGAGCGCGCGAATACGCGACGCTTCTTGACATATCTACGAAACGACAGAG	274099
Oy	213	AlaAaenProArgMetLeuIleValHisGluValMetGlyArgAaBnCyBglYTrpLeuThr	232
Db	274100	GCCGCGCGCGGACCCCTCGCTCGTCAATGAAGTCATGGCGCGACACTCGCGCTGGTGAAC	274159
Oy	233	AlaAlaThrAlaGlnGlnIleuTyArgIleuLeuAaBpArgAlaGluTrpLeuProGlnLeu	252
Db	274160	GCGGACACCGCGCGCTATATTCATCTGCGCGGCAACAGAAATATCTCGACGGCTTC	274219
Oy	253	GlyLeuThrArgGlnSerTyArgIleuValHisAlaValPheAlaProIleMetAlaIleAaP	272
Db	274220	ATGATGAACGACAGCTGAAGAACATCGACGGCTTACTCTTCGAGATGGGTTTCGAT	274279
Oy	273	LeuGlnAlaGluAlaIleuAaBpArgIleuValMetAaBpIleValAaBpCyValAaBnIle	292
Db	274280	CTCGAAGCGGAAGCGCGCGCTCGCGCAAGTCATGATCGAAGCGGCTTCGACGTTG	274339
Oy	293	PheValSerGlnGlyValaGlyValaGluAlaIleValAlaGlnMetClnAlaIleuBglGln	312
Db	274340	TTCTGAGAGGAAGGGCGCTCGCTCGACGCTCTCGCGAAGCGGAGCGCGCGCGAG	274399
Oy	313	GluValProArgAaBpAlaPheGlyHisIleuLeuAaBpAlaValAaBpProGlyIleuTrp	332
Db	274400	ACGGTCACAGCGGACGCGCTTCGCGCATGTGAAGATGATGATCAACCTCGGAACCTGG	274459
Oy	333	PheGlyGlnGlnPheAlaGlnMetIleGlyAlaGluIleuThrLeuValGlnIleuSerGly	352
Db	274460	TTCTCCAGCAATGTGCGCGCGCTCCCTCGCGCGCGACGCTCATGTGTGCAAGATCGGC	274519
Oy	353	TyrPheAlaArgAlaSerAlaSerAaBnValAaBpMetArgIleuIleuSerCyAla	372
Db	274520	TACTATGACGCTCGCGCGCGCCCAATGTGCAAGATCGCTCATCCAGACATGATC	274579
Oy	373	AaPheAlaValaGlnCyAaBpAaBpAaBpArgIleuSerGlyValIleGlyHisAaBpGluAaP	392
Db	274580	GATCTGGCGGTGAGAGCGCGCTCAACAAGATCTCCGGCGTGAAGCGGCCACAGAGAT	274639
Oy	393	AaBglAaBnValIleuAaBpAlaIleGlnPheProArgIleuBglGlyIleuBpProPheAa	412
Db	274640	CAGGCGCGGCGCTGCGACATCGATCGATTCCGCGCATCAAGGCGGAAGCATTTCCGAC	274699
Oy	413	IleAaPThrAaPTrpPheAaBnSerMetLeuSerGlnIleGlyGln	427
Db	274700	ACGTCCGCGAAGTGTTCGGCGAGCATGATGTCTGCGCGAG	274744
RESULT 13			
AE008127	11715 bp	DNA	linear BCT 28-MAY-2004
LOCUS			
DEFINITION	Agrobacterium tumefaciens str. C58 circular chromosome, section 185 of 254 of the complete sequence.		
ACCESSION	AE008127	AE007869	
VERSION	AE008127.1	GI:15157238	
KEYWORDS	Agrobacterium tumefaciens str. C58		
SOURCE	Agrobacterium tumefaciens str. C58		
ORGANISM	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.		

REFERENCE	JOURNAL	FEATURES
1. (bases 1 to 11715) Hinkle,G., Slater,S.C. and Goodner,B. Complete Genome Sequence of Agrobacterium tumefaciens C58 (Rizobium radiobacter C58), the Causative Agent of Crown Gall Disease in Plants	unpublished	source
2. (bases 1 to 11715) Hinkle,G., Slater,S.C. and Goodner,B. Direct Submission Submitted (14-AUG-2001) Cereon Genomics, 45 Sidney Street, Cambridge, MA 02139, USA	unpublished	location/Qualifiers
		1. 11715
		/organism="Agrobacterium tumefaciens str. C58"
		/mol_type="genomic DNA"
		/strain="C58"
		/isolate="Cereon"
		/db_xref="taxon:176299"
		289 . 693
		/gene="AGR_C_3816"
		289 . 693
		/gene="AGR_C_3816"
		/note="HYPOTHEICAL 14.2 KD PROTEIN IN VACB-AIDB INTERGENIC REGION (0132)"
		/codon_start=1
		/transl_table=11
		/product="AGR_C_3816p"
		/protein_id="AAK87853.1"
		/db_xref="GI:15157239"
		/translation="MMTIVAAAGLPAFLSYSGVCPAFSTIYTRLTPHEVELIKAA GNLSAVAFICGLIGFSLPLASAAHSVSLPDVITPMADYGLVQLAFYAKLTMDL HKLTENVAAGLMSGGIAVYIGTLNACMAY"
		703 . 1299
		/gene="AGR_C_3818"
		703 . 1299
		/gene="AGR_C_3818"
		/note="hypothetical protein"
		/codon_start=1
		/transl_table=11
		/product="AGR_C_3818p"
		/protein_id="AAK87854.1"
		/db_xref="GI:15157240"
		/translation="MHLDVTEHPRALAGTIAIGLFPASTEEBLYTSDVKACASGV YPRACQASEADANQHLATAPRFNNVACMEYGAAGCAOLROSVDNSSTSAFVY SPGFWPVRVENFSDSEYRRQNNSSGASAMSAVRRNGDLVTPDVAQQNNNG GGGAGAGGALADVDTLNAKTNTVKGRCGCHSFGSG"
		1301 . 1846
		/gene="AGR_C_3820"
		1301 . 1846
		/gene="AGR_C_3820"
		/note="HYPOTHEICAL 45.0 KD PROTEIN IN TOLC-RIBB INTERGENIC REGION (0386)"
		/codon_start=1
		/transl_table=11
		/product="AGR_C_3820p"
		/protein_id="AAK87855.1"
		/db_xref="GI:15157241"
		/translation="MRRITVPPRPDRMEKANAVGSEHNGEBYVDDAAYCGLEE IENRIEBSQILDMCDMLVADIVNSEALTRLAIPTADYDVTINSMRGDRHL YGRF DLAYDGRPAKLEYNADPTSTVFETCYFOYNNMLTDVALGTLPADTDQNSVQEAII EAREQFSRDMGRKISWTAL"
		complement (1813 . 3419)
		/note="IS1353-1like insertion element"
		complement (1841 . 3379)
		/gene="AGR_C_3822"
		complement (1841 . 3379)
		/gene="AGR_C_3822"
		/note="possible transposase of IS1353"
		/codon_start=1
		/transl_table=11
		/product="AGR_C_3822p"
		/protein_id="AAK87856.1"
		/db_xref="GI:15157242"

```

gene
3439..4068
/ gene="AGR_C_3823"
/ protein_id="AAK87857.1"
/ db_xref="GI:15157243"
/ note="HYPOTHEICAL 45.0 KD PROTEIN IN TOLC-RIBB INTERGENIC REGION (ORFA) (O386)"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3823p"
/ protein_id="AAK87857.1"
/ db_xref="GI:15157243"
/ translation="MSDSEDGCTTLYLMDCAIOGHRAPFDIRDIGDASGRVADL OSRVIERCKLYPMEFMLEPESRANLPAESGPIRPAWALLSNGLPMLQRRPHN PNLPSFPDDPEIYELQDYVYKPLSRGENVSLFRNREILSPAGSTGKGYVFOA YAPLPSFEGYAVLGSWIVGDRSCGLIREVDVSPITANLSRYVPHIEG"
/ complement(4553..5794)
/ gene="AGR_C_3826"
/ complement(4553..5794)
/ note="SARCOSINE OXIDASE"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3826p"
/ protein_id="AAK87858.1"
/ db_xref="GI:15157244"
/ translation="MQATRMKRVNGTCKEGSGHPNTYVGRGMGAANAHLAETV DGVALLIGEPADIKSHQGVFASHYDEARITITIGDDPMWLLANRSIARYADIDARS GVEFYAPVOCMLWGPBEGANFPVDVLAARLIGVSTELLGDOSLKSFPYFSEEPG CEGFEPBDNAGVYNPRALVKAQALIAEKAGVTLIDIVYSTRDEGRASVOTASGVV TAERYLVNAGSGFSTRDLIPQVLDVARTVAPRIDEGLDGOYAGMPSLIYEPDHT TKHYLLPVRYPDQKFTYIKGCGDDPDKRQVSDPEIRERFSGGREGSRVDSYVGT LVPYSDHSVSMACVSKTKSGYPALGFTASPRIAVALTGNGTAAKSSDEIGRLGAA LRDKIDADDASTDQKPEFL"
/ complement(5803..7116)
/ gene="AGR_C_3830"
/ complement(5803..7116)
/ note="Chain A, Complex Of Monomeric Sarcosine Oxidase With The Inhibitor Dimethylglycine"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3830p"
/ protein_id="AAK87859.1"
/ db_xref="GI:15157245"
/ translation="MROAGIAGELRPASIIICEIAKNDGSGRHAHVNRKSDTEHMP DEXYLVGAGMGAARHLISYOTGVALLIGAPADRKTHQGVSSHVNRVIRTRGF DGDPMVLAELASIRRYAIEBAKSGIRFTEGCLFTGNGKLAGDVYSNADRL GLGVETIAGAEALAGRPMFALPADHSGSEANACHINRALVYKQCAIAEAGCGRV RETAAHIRTDSRSEVMTREGATYTAEKITVAAAGTNNALLPSPVDMAAGRTIVF FELDARQALFSAMPSTIVLAEITDVIYLPVRPDQKVLTKIGSEKRLSTLV QAVDFHSDGTDEVEFLTKKALSLMPALAGCPVSGSCVASITSGSYIYTOSSN IAVLTGNSVNSAKSDEIGRLGALLNQLTEDEFAEMSVFV"
/ complement(7083..7670)
/ gene="AGR_C_3831"
/ complement(7083..7670)
/ note="putative transglycosylase"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3831p"
/ protein_id="AAK87860.1"
/ db_xref="GI:15157246"
/ translation="MPFHCCHPIWRILFLHNRFPVAVVAVGVALASQAGISPAKDED

```

```

gene
VKTQVTLLETVEVRKPGYPIPEKSLPASLKNYSDLIVTKAKRYGPTNLAAHVSVES KENPARAGSAGVGMQIKPATARMGFGATKALYDPETNIRMGQYDATALQLOGG EVCSITLRNAGAGATRMNVPKRCVKQVALLAS"
/ complement(7870..8649)
/ gene="AGR_C_3833"
/ complement(7870..8649)
/ note="conserved hypothetical protein PA5485 (imported) - Pseudomonas aeruginosa (strain PAO1)"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3833p"
/ protein_id="AAK87861.1"
/ db_xref="GI:15157247"
/ translation="MKECIPDPAFDPFGCAVAPSPNHGRLGVAGPDIIILLHYTGWT ADGALSMWRNPSOVSSHVFEDRGRVQIOLVPSRRAMAGASVAGDEIDNSRIGI EIANQHGPGGLPEPEAOVAAVIELCRDQGRWMSAPREVLAHSDVAPTRKYDPEKFP MDILSQHGHWVEPAPTRGRFFROGDHQPVEALQSMLSITYGAEITGAICYCKT EGAVVAFOHFRPRLVDGIADFTITDTHRLISSLPFSAA"
/ complement(8646..9422)
/ gene="AGR_C_3834"
/ complement(8646..9422)
/ note="(AB005182) putative DNA binding protein"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3834p"
/ protein_id="AAK87862.1"
/ db_xref="GI:15157248"
/ translation="MHSREYPPMLASMPDFPCLQISLMEWLLAIGDAGNALGRV VEARITLPEGDPETRRKVSFVAIILASAKMAKAGVNDADAVRAFROI FDPDEBAK NVARLYNLARQOVAGYEAVERIAGLCSGHANCMLEISVIGLFTIAADGLIHRE IAFIRGIAIFRITDEHPETIMARIYHMDGRDPRVAGVSPDDFLDIRKRYSLVAB HHPDLIARGVVMEHAAANERMAALNAVAIAEKERRVA"
/ complement(9385..10929)
/ gene="AGR_C_3836"
/ complement(9385..10929)
/ note="(AF246209) pPl-phosphofructokinase"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3836p"
/ protein_id="AAK87863.1"
/ db_xref="GI:15157249"

```

Alignment Scores:

Pred. No.:	Length:	Matches:	Mismatches:	Indels:	Gaps:
3,26e-86	11715	246	71	101	7
Score:	1292.50	Matched:	Conservative:		
Percent Similarity:	74.59%	57.88%			
Best Local Similarity:	56.59%				
Query Match:	1				
DB:	1				

```

US-10-701-200-6 (1-437) x AEO08127 (1-11715)
Qy 8 HsLeuThAlAspIle-----ArgpheCyHisTrpPheheLeuAenphe 23
    |||||  |||  ::  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 9649 CACCTTACACGCGGCTCTCCATCCATTCGCCAGACATGCTTGGGTATACGCAAAA---- 9705
Qy 24 AenpheTyTrhLeuWecAenlysProlysValAlaIleLeuThAlaIglyIleu 43
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 9706 -----AGGAGAGCATTATGCGAAGAACAGAAAGTCGCAATGCTGACCCGGGTGGCCTC 9759
Qy 44 AaPrcCyLeuAaSerAaIleGlySerLeuIleGluAaTyTrhGluIleAaPrc 63
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 9760 GGGCCCTGTCTCTTCGGCGCGCTGATCGAGCGGCTATAGCGACATCGGCGCT 9819
Qy 64 SerIleGluIleIleCyTyArgGlyIleTyGlyLeuLeuLeuGlyAaPserTy 83
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 9820 GAAATGACATCATCGCTTACCGCTCCGTTATCAGGCGGTCTTGGCGACGCTATC 9879
Qy 84 ProvalThralGluValArgLysLysAlaGlyValLeuGluAlaArgPheGlySerVal 103
    ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||

```

Db 9880 GAGATCACCAGAGATATGCGCGCAAAAGGCGCATCTACATCGCTTACGGCGCTCGCCG 9939
 Qy 104 IIGGIAAASerAryValIlyLeuThraenValIyAAspCyAValIyAArgIyLeuVal 123
 Db 9940 ATCGGCAACAGCGCGGTCAAGTCAACATGCGCGCATTTGCGCAAGCGCGCTCGTC 9999
 Qy 124 IyAGIyGlyIyAAspProGlnIyValIAlaIAspGlnIyValIyAAspGlyValAAsp 143
 Db 10000 AAGGAAGCGCAATCCGTGGCGCTCGCGCTGGAACGGCTGCGCGCTGACCGCATCAC 10059
 Qy 144 ILeuAHisThraIIGIyValIyAAspThraenThraIAlaIAspLeuIAlaIAsp 163
 Db 10060 ATTCTCCATCTATGCTGGTGGCCACACACCAACACCGCGCGCATTTCCCGCTTAT 10119
 Qy 164 ILeuAAspAAspAAspIyLeuThraIIGIyLeuProIyThraValIAspAAsp 183
 Db 10120 CTGGCGCGCAACGATATGATCACTTCCGCGCTGCGCGCTGCAAGCGGTGATATATGAC 10179
 Qy 184 ValPheProIleIyGlnIySerIyLeuIyAATPTThraIAlaIAGIyGlnIyValAArgIy 203
 Db 10180 GTCGTGCGCATACAGCTGCTGCGCGCTGCGCGCTGCGCGCGGAGTCTGCTGCTTC 10239
 Qy 204 PheMetAAspValIAlaIyAAspAAspAAspProAArgMetLeuIleValIAGIyVal 223
 Db 10240 TTGCAATATGTACGACAGCAGCAGCGCGCGCAACCGCAACCTTGCTATCCACAGATC 10239
 Qy 224 MetGlyAAspAAspGlyIyTTPleuThraIAlaIAThraIAGIyIyAArgIyLeuLeu 243
 Db 10300 ATGGCGCGCATGCGCGCTGCGCTGACCGCGCGCGCGCGCTATATCCAGAAACC 10359
 Qy 244 AAspAAGIAGIyTTPleuProGlnIyLeuIyAATThraIArgIyIySerIyGlyValIAla 263
 Db 10360 CGCGGCAACGATATGCTGCAAGCGCTGATGATGAACACCGCAATGAAGAACATCAACGCGC 10419
 Qy 264 ValPheValIProGlnIyMetAlaIleAAspLeuGlnIyAGIyValAArgIyLeuArgIyVal 283
 Db 10420 ATCTACCTCGCGCAATGCTCTTCAATCGACGCGCGCAAGCGCGCTCAAAAGATC 10479
 Qy 284 MetAAspIyValAAspCyAValAAspIlePheValIySerGlnIyAGIyValIAGIyVal 303
 Db 10480 ATGGAACAAACATGCTATGTACGCTGCTGCTTCCGAAAGTGGCGCGCTCGACCGCATC 10539
 Qy 304 ValAlaIAGIyMetGlnIyAAspIyGlnIyValIProAArgAAspAlaPheGlyIyIleIy 323
 Db 10540 GTCGCGCGCAAGCGCGCTGCTGCGCGCAAGCGCGTCAAGCGCGATGCTTCCGCGCATGGAAG 10599
 Qy 324 ILeuAAspAlaValAAspProGlnIyTTPheIyGlnIyGlnIyPheAlaIAGIyMetIleGlyAla 343
 Db 10600 ATCGACACCATCATGTTGGCGGTGCTTCCAGAAAGATTCGCGCGCTCATCGCGCGC 10659
 Qy 344 GlnIyThraIyValIAGIyAAspIySerGlyIyTTPheAlaIArgAlaIySerAAspValAAsp 363
 Db 10660 GAGCGCTCATGATGTGAGAAAGTGGCGCTATTTCCGCGCGCTCGCGCGCAATGGCGAGC 10719
 Qy 364 AAspMetAAspLeuIleIyAAspIySerCyAAspLeuIAlaIAGIyValAAspAArgIyGlu 383
 Db 10720 GATCTTCGCTCATTCACGCGCATGTGCTGCGCGTGAAGCGCGCTCATATAGATC 10779
 Qy 384 SerGlyValIleIyGlnIyAAspGlnIyAAspAAspAAspValIyAArgAlaIleGlnIyPhePro 403
 Db 10780 TCCGCGGTACCGCGTATATCAAGACCAAGCAATGGCAATTCGCGACCATTCAGATTTCG 10839
 Qy 404 ArgIleIyGlyIyValIyAAspProPheAAspIleAAspThraIyTTPheAAspMetLeuSer 423
 Db 10840 CCTATCAAGCGCGCAAGCATTTCCGACTCTCTGCAAAATGTTTACCGAAGATGAGAA 10899
 Qy 424 GlnIleIyGlnIyPro 428
 Db 10900 CACGTGCTGACGCG 10914
 RESULT 14
 AEO09161
 LOCUS AEO09161 11807 bp DNA linear BCT 28-MAY-2004

DEFINITION Agrobacterium tumefaciens str. C58 circular chromosome, section 187
 of 256 of the complete sequence.
 ACCESSION AEO09161 AEO08688
 VERSION AEO09161.1 GI:17740565
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 The genome of the natural genetic engineer Agrobacterium tumefaciens C58
 JOURNAL Science 294 (5550), 2317-2323 (2001)
 MEDLINE 21608550
 PUBMED 11743193
 REFERENCES
 AUTHORS
 2 (bases 1 to 11807)
 Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Chen, Y., Woo, L., Kiteajima, J.P., Okura, V.K., Almeida Jr., N.F., Zhou, Y., Boye Sr., D., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Guenther, D., Kutayvin, T., Levy, R., Li, M., McClelland, E., Palmer, A., Raymond, C., Rouse, G., Saenphimachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V. and Nester, E.W.
 Direct Submission
 Submitted (27-SEP-2001) Department of Microbiology, University of Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA 98195-7242, USA
 TITLE
 JOURNAL
 FEATURES
 source
 1..11807
 /organism="Agrobacterium tumefaciens str. C58"
 /mol_type="genomic DNA"
 /strain="C58"
 /isolate="U. Washington"
 /db_xref="taxon:176299"
 384..785
 /gene="Atu2104"
 384..785
 /gene="Atu2104"
 /note="identified by sequence similarity; putative; ORF located using BLaetx/Glimmer"
 /codon_start=1
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="AAL43095.1"
 /db_xref="GI:17740566"
 /translation="MTIYAAGAPLAFSYGIVCFALFSTIYTRTPHEVELIKAG NISAVAFILGILGFSILPLASAAHVSFLPDVIVAAIGLVQVLAFLAKLTWMDLH LKIRGNVAAGIMSGIIVIGTINAAACMA"
 795..1391
 /gene="Atu2105"
 795..1391
 /gene="Atu2105"
 /note="putative; ORF located using Glimmer"
 /codon_start=1
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="AAL43096.1"
 /db_xref="GI:17740567"
 /translation="WHLDVTEPRPLALGTIAIGLPMASFTSEBVLVYTSVDKVASGV YPKAQASSEDANQHHLATPRFRMVACSEHYGANGCAQLQSVNSDSSSTSAFVP

gene
SDGAAGMPMSVNFPSDVFYSRRROONSSNGSAGAAAYARKKXGDLVTIPDVASGQNNNG
GGGAGGAGGALNDPDTLNAKTNVAGRGFGHSPSGS"
1393..1398
/gene="grp"
/note="synonym: Atu2106"
1393..1398
/gene="grp"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="glutathionylspermidine synthetase (N-terminal)"
/protein_id="AAL43097.1"
/db_xref="GI:17740569"
/translation="MRRTVPDPDKRANAVGSFHENGEPYWDADAYCGLEE
IRNLIEEPPQILHDMDLVADIVSEBALTRLAIPLYOHVISNRMGGRHLVGRFP
DLAYHGPGQLLEVNADPTPTSFTGYFOYMTLDQVALGTLPADTDFFNSVOEALII
BAFGFSRDRLMVGRIKSXTAL"
complement(1933..3471)
/gene="tnp"
/note="synonym: Atu2107"
complement(1933..3471)
/gene="tnp"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="IS3 family transposase"
/protein_id="AAL43098.1"
/db_xref="GI:17740569"
/translation="MVSVDRLRAVELYIRLGRLLNTTITROLGYPTKNALRGWREVV
QHLDRTQPVAAPKYSEARQALAEHRTRDRCSGTMRALGTPGREGTLTAWRERAP
PEARSTWGRSNHPGVSEVRQAGVIGLCSDGESAQOVDLGSRPFLYSKKQGLLG
HAASSMKRKANPKVPEREEELERLELQORVROLQEHDLKKAMELKKGGLVD
LILNSEKQILDALKEVYRLPELLAQCLIARSYFYHARVMCLADKAAATLSLABI
FEANRCYKTIOLQSLARKSVYSEKVORLMKOEHVAARPERRRSGVLGISPA
PNLTINRDHAAPNVKMLTDITEROIAGKYVLSPIIDCPDGVISSIGTOPDAAL
VNTMLDAALGTVAANGERRITHSDRGAHYPRPGMLTRISEKLVSNSRKCSDNNAI
CGGFPERLTELFPDMKAITTEQVAEVDAIYRWNEKRKIISLQSLSPVEYRQSL
GIKL"

gene
3564..4160
/gene="grp"
/note="synonym: Atu2108"
3564..4160
/gene="grp"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="glutathionylspermidine synthetase (C-terminal)"
/protein_id="AAL43099.1"
/db_xref="GI:17740570"
/translation="MYLMDCAIQAGRAQPLDIRDIGIDASGRVADLOSRYTERCPFKL
VYMEMLRPSFPANLPAESGFIPEAWRALISNKGLPMIOGNPHNHPLLPSPPDD
PEYIELQDIVVKPLRSBGENSVLYRNRETLSAFGSYGQVAYALPFSEFGCT
AVLGSWTIVGSRGCGIREDVSPITANLSRYPHILEG"
complement(4188..4589)
/gene="Atu2109"
complement(4188.
/gene="Atu2109"
complement(4188..4589)
/product="RNAase P RNA (tmprb) "
/note="Found by BLASTn"
complement(4645..5835)
/gene="solA"
/note="synonym: Atu2110"
complement(4645..5835)
/gene="solA"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="sarcosine oxidase"
/protein_id="AAL43100.1"

CDS

gene
miac_RNA

gene
CDS

gene	complement (5895 . 7082)	
CDS	/translation="MSGHNYIVVGRGMMGAAARHLETVDVGLIGPSPADIKSH /db_xref="GI:17740571" /feature="gltc" /note="synonym: Actu2111" complement (5895 . 7082) /gene="gola" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer" /codon_start=1 /transl_table=1 /product="sarcosine oxidase" /protein_id="AA143101.1" /db_xref="GI:17740572" /translation="MPDFKYIVVAGAGMGAARHLSVQTDGVALIGPSPADIKTHQ GVSSHYDVARTIRGPDGDPVMAELARSIRVAELBAKSGIRPFTAGCLFTNGNGK LAGYVSRALSSADRLGVELTGAEALAGRPFMPALPDHGSFEFNRNGHILPAPL VKQACALAEAGQGLVREETAHRIQDTSRVSVMREGATYTAEKVIVAAGFTMALV LSPVDMAATGRTIVPELDEARQALFSAMPSTVLAETEDDIYIILPVRYPGKXYV LKTGSESEKRELTIVQADVWFHSDGDFVEFTTKALSLMPALACPVTSGCVAS ITGSGPYIIGYTOSSNIAVLVTGNGFVSAKSDIEGRGLAOLLINGLTDEFEAEMSEB VPV"	
gene	complement (7175 . 7714)	
CDS	/gene="gltc" /note="synonym: Actu2112" complement (7175 . 7714) /gene="gltc" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer" /codon_start=1 /transl_table=1 /product="soluble lytic transglycosylase" /protein_id="AA143102.1" /db_xref="GI:17740573" /translation="MNRFEVAVVAVGVALASQAGISPAKDEYTKTIVTLETPRKSG YPIPEKLPASIKNGSDLIYKAKRGVGTNTLAHVAISVSKNPAICAPVSGVGLM QIRPARKMGFRGATRALYDPEPTNIRMGQYLAATHQLGGGVCSTILKRNAGHAT RMPVSKRYCGKVQALAS"	
gene	complement (7962 . 8741)	
CDS	/gene="Actu2113" complement (7962 . 8741) /gene="Actu2113" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer" /codon_start=1 /transl_table=1	
Alignment Scores:		
Pred. No.:	3.29e-86	Length: 11807
Score:	1292.50	Matches: 246
Percent Similarity:	74.59%	Conservative: 71
Best local Similarity:	57.88%	Mismatches: 101
Query Match:	56.59%	Indels: 7
DB:	1	Gaps: 2
US-10-701-200-6 (1-437) x AE009161 (1-11807)		
OY	8 Histenuhralaaple-----ArphecyshtTpphepLeuanpHe 23	
Db	9741 CACCTTACACGGCGGTGTCGATCCATTTGCCAGACATGCTTTGGGATACGCAAAA--- 9797	
OY	24 AsnpheryrthleuMeastnlyspIolylysvaIaaleuNrtaIaglyIeu 43	
Db	9798 -----AGAGAGCGCATGATGCGCAAAACAGAAATCCCATGTCGACCGCGGTCCTC 9851	
OY	44 AlaprocysleuanseRralaIleglySerleuilegluarGlyrthrgluIlaapPro 63	

```

Db      9852 GCGCCCTGCTCTCTCGCCCGGTGGATGGCCGATCGAGCCGCTAATGACATCCGCT 9911
Qy      64 SerilegilelleleCyetyrArglygylYrlysglyLeuleuLeuglyApseryr 83
Db      9912 GAATCGACATATGCTTACCGCTCCGGTTATCGAGGCGTCTCTGGCGACAGCTATC 9971
Qy      84 ProvalThrAlaGluValArglyLeuValaglyValLeuGlnArgPheglyYseryVal 103
Db      9972 GAGATACCAAGATATGCGCAAAAGGCGCATCTACATGCACTGTTACCGCGCTCGCG 10031
Qy      104 lileglaanserArgvallyleuThrAenVallysaPcyVallyAarglylyleuVal 123
Db      10032 ATCGGCAACAGCGCGGTCAAGCTGACCAATGCCCGCATTTGCCAAGCGCGCTCGTC 10091
Qy      124 lylsglygllyAapProGlnlyValAlaAlaAapGlnleuVallysaPglyValAap 143
Db      10092 AAGGAAGCGCAATCCGTGGCGCTCGCGCTGACAGCGCTGGCGGCTGACCGCATCAC 10151
Qy      144 lileuHsthrileglyyAapAapThrAenThrAlaAlaAapleuAlaAaphe 163
Db      10152 ATTCCTCATCTATCGGTGCGACGACCAACACACCGCGCGGACCTTCCGCTAT 10211
Qy      164 LeuAlaAapAenThrlyleuThrValileglyleuProlyleuThrAlaAapAap 183
Db      10212 CTGGCGCCCAAGATATGATACCTCCTCGCTGCGCTGCGAAGACGCTGATATGAC 10271
Qy      184 ValPheProilelelyleuSerleuGlyAlaATPThrAlaAlaGlyGlnlyAlaArglyr 203
Db      10272 GTCGTCGCGCATACAGACGTCGCTGGCGCTGACGCGCGCGGAGTCCGCTCTTC 10331
Qy      204 PheMetAenValAlaAlaGlnAenAalAapProArgMetleuileValHieglyVal 223
Db      10332 TTCGCAATGTGAGCAACGAGAGCGCGGACGCAAAACCTTGCTATCAGCAAGTC 10391
Qy      224 MetGlyAargAenCyglyYrlyleuThrAlaAlaATPThrAlaGlnlyYrlyleuLeu 243
Db      10392 ATGGGCGCGCATTTGGCTGGCTGACCGCGCGCGCGCGCTATATCCAGAAACC 10451
Qy      244 AapArgAlaGlnlyleuProGlnlyleuThrArgGlyserYrlyleuAlaHisa 263
Db      10452 CGCGGCAACGATATGTCGAAGGCTGATGATGAAACGCAATGAAGAATCCAGCGC 10511
Qy      264 ValPheValProGlnMetAlaileleleuGlnlyAlaAlaArglyleuArglyVal 283
Db      10512 ATCTACCTCCCGAAATGCTCTTCGACATCGAGCGGAGCGGCTCAAGAGGCTC 10571
Qy      284 MetAaplyValAapCyValAenilePheValserGlyAlaGlyValAlaile 303
Db      10572 ATGGCAAAACATGCTATGTCAGCTGTCGTTCCGAAGGTGCGGCGCTCGACGCGCATC 10631
Qy      304 ValAlaGlnMetGlnAlaAlaArglyGlnlyValProArgAapAlaPheglyHieileys 323
Db      10632 GTCGCGGAGCGGGAAGCTCTGCGCAAGCGCTCAAGCGGATGCTTCGCGCATGGAAG 10691
Qy      324 LeuAapAlaValAapProGlyYrlyrPheglyGlnlyGlnPheAlaGlnMetilegylA 343
Db      10692 ATCGACACCATCAATGTTGGCGCTGTTCCAGAGCAATTCGCGCGCTCATCGCGCGC 10751
Qy      344 GlnlyrThrleuValGlnlyserGlyYrlyrPheAlaArgAlaSerAlaSerAenValAap 363
Db      10752 GAGCGCTCATGTCGTCGAAGGTGCGGCTATTTCCGCGCGCTCGCGCGCATGCGGAC 10811
Qy      364 AapMetArgleuileyseryCyAlaAapleuAlaAlaGlyCyAlaAapheArgly 383
Db      10812 GATCTTCGCTCATCGAGCGCATGTCATCTGCGCGCTGAAAGCGCGCTCATAGATC 10871
Qy      384 SerGlyValileglyYrlysaPglAapAenGlyAenValleuArgAlaileGlyuphePro 403
Db      10872 TCCGGGGTCAACCGGTATGACGAGACGAAATGCGAAATTCGCGGACATCGAGTTCCG 10931
Qy      404 ArgGlyleuGlyYrlysaPheAenileAapThrAapThrAapSerMetleuSer 423

```

```

Db      10932 CGTATCAAGGCGGAAAGCATTTGACCTCTCTGCAAAATGTTTACCGAAGTATGAA 10991
Qy      424 GlnileglyGlnPro 428
Db      10992 CACGTGCTCAGCGC 11006

RESULT 15
SCU51728
LOCUS SCU51728
DEFINITION Streptomyces coelicolor phosphofructokinase ATP-PFK gene, complete cds.
ACCESSION U51728
VERSION U51728.1 GI:1931572
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCES
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1549
/organism="Streptomyces coelicolor"
/mol_type="genomic DNA"
/strain="1109"
/db_xref="taxon:1902"
160..1188
/codon_start=1
/transl_table=1
/product="phosphofructokinase ATP-PFK"
/protein_id="AAC45135.1"
/db_xref="GI:1931573"
/translation="MKVGVLTGGGDPGLNATVAVKVGQVGYDFTFGRDGRGP
LEGDTVPDIPAVRGILPRGGVILSSSTNPLKQDGRIRIKDNLAALGVALITIGG
EDTGVATRIADDEGVPCVGPXTIDNLSATDYPFGDTAVGATBIAIDRLHTAES
HMRVIVYVMGRHGMIMLHSGIAGGAAVILIPEROVEOVCSVTSRPSVAPIV
VVARGAMPRDDDMVKDSLPSYGHVRLSGVEMLAKQIERKTGNEARTTVLGVORG
GTPSAFDRWLATRGLHAVDCVHDDPFGKVALRGTDIVRPIATATRLKTVDPALY
EEVGVFPG"

ORIGIN
Alignment Scores:
Pred. No.: 5.56e-14 Length: 1549
Score: 311.00 Matches: 105
Percent Similarity: 43.86% Conservative: 63
Best Local Similarity: 27.42% Mismatches: 137
Query Match: 13.62% Indels: 78
Gaps: 17

US-10-701-200-6 (1-437) x SCU51728 (1-1549)
Qy      34 LysValAlaileleuThrAlaGlyGlyleuAlaProCyApseryrAlailegysery 53
Db      163 AAGGTGAGATCTACGAGGAGCGGCGGACATGCTCCCGGCTCAAGCGCTATCCGGCGC 222
Qy      54 LeuileglaArglyrThrGlnleuAapProserilelelleCyetyrArglygyl 73
Db      223 GTCGTCGCGAAGGCGGTCCAG---GAGTACGCGCTACGACTTACCGGCTTCCGGAGCGC 279
Qy      74 TyrlysglyleuLeuLeuglyAapserYrProValThrAlaGlnlyVallysa 93

```

Job time : 6152 secs

```
Db 280 TGGCGGGCCCCCTGGAGGCGCACCGCTCCCGCTC-----GACATCCCGGGCGCTCCGC 333
Qy 94 GlyValIleuGlnArgPheGlyGlySerValIleGlyAsnSerArgValIleLeuThrAsn 113
Db 334 GGCATCTCTGCCCCGCG---GGCGGCACCGCTCTGCGCTCTCTCCGCAACCGCGCTCAAG 390
Qy 114 ValIleAspCysValIleArgGlyLeuValIleGlyGlyLeuAspProGlnIleValIle 133
Db 391 CAGCGGGACGGCATCCGGCGC-----ATC 414
Qy 134 AlaAspGlnLeuValIleAspGlyValAspIleLeuHisThrIleGlyGlyAspThr 153
Db 415 AAGGACACCTCCCGCACCTCCGCGGCTCATCCATCCGCGCGGAGACACC 474
Qy 154 AsnThrAlaAlaIleAspLeuAlaIlePheLeuAlaArgAsnAsnThrGlyLeuThrVal 173
Db 475 CTGCGGCGTGGCCACCGCGCTCCGC-----GACGAGTACGGCGTGCCTGC 519
Qy 174 IleGlyLeuProIleThrValAspAsnAspValPheProIleGlyGlnSerLeuGlyAla 193
Db 520 GTCGGCGTCCCGCAAGACATCGACAGACCTGCGCCACGACCTACACCTTCGGGTTTC 579
Qy 194 TrpThrAlaIleGlnGlnGlyAlaArgIlePheMetAsnValValIleGluAsnAsnAla 213
Db 580 GACACCGCGCTCGCATCCGACCGAGCCATCGACCGCTGCACACACCGCGAGTCC 639
Qy 214 AsnProArgMetLeuIleValIleGlyValIleMetGlyArgAsnGlyGlyTrpLeuThrAla 233
Db 640 CACATCGGCTCTGGTGGTC---GAGGTATGCGCGGACCGCGCTGATAGCCCTC 696
Qy 234 AlaThrAlaGlnGlnIleTrpArgIleLeuAspArgAlaGlnTrpLeuProGlnIleGly 253
Db 697 CACTCG-----GGC 705
Qy 254 LeuThrArgGlnSerTrpGluValIleAlaValIlePheValProGlnMetAlaIleAspLeu 273
Db 706 CTGGCCGCGGCGCGC-----AACGTATCTCATCCCGACGACGCGCTTCGACGTC 756
Qy 274 GluAla-----GluAlaIleArgLeuArgGluValIleMetAspIleValAspCys 289
Db 757 GACGAGGTGCTCTCGGTGACCTCCGCTTCGGGCGCTCTACGCGCCGATC----- 810
Qy 290 ValAsnIlePheValSerGlnGlyAlaGlyValIleGluAlaIleValIleGlnMetGlnAla 309
Db 811 -----GTGGTCTCGCGGCGCGC-----ATGCCGCGC 840
Qy 310 LysGlyGlnGluValProArg-----AspAlaPheGlyHisIleIleValAsp 325
Db 841 GACGCGGACATGTGCTCAAGACGAGTCGCTGACCTCTCACTCAAGGACGTCGGTCTCC 900
Qy 326 AlaValAsnProGlnIleTrpPheGlyGlnGlnPheAlaGlnMetIleGlyAlaGlu--- 344
Db 901 GGGGTC-----GGCAATGGCTGGCCAGACATCGAGAGCGCACCGCAAGAGCC 954
Qy 345 LysThrLeuValGlnIleSerGlyTrpPheAlaArgAlaSerAlaSerAsnValAspAsp 364
Db 955 CGACACCACTGTC-----CTCGGCAAGCTCCAGCGCGGCGCACCGCCGCTTCGAC 1008
Qy 365 MetArgLeuIleLysSerCysAlaAspLeuAlaValIleGluCysAlaPheArgArgIleSer 384
Db 1009 CGCTGCTCGCCACCGCTTCGAGCTGACAGCGCTCGACTGCGTG----- 1053
Qy 385 GlyValIleGlyHisAspGluAspAsnGlyAsnVal-----LeuArgAlaIleGluPhe 402
Db 1054 -----CACGACGGGACTTCGCAAGATGTGCTCGCGGCGACGACATC 1101
Qy 403 ProArgIle 405
Db 1102 GTCGCGCTC 1110
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2005, 01:38:40 / Search time 721 Seconds
(without alignments)
3587.971 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVATWPHYLTADIRCFCHMF.....FNSMLSEIQPKGKRVESH 437

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODBI=framer+ p2n model -DRV=hlh
-Q=/cgn2.1/USPTO.spool/US10701200/runat.25022005.102832.10300/app.query.fasta_1.583
-DB=N.Geneseq.16Dec04 -OPMT=fastap -SUFITX=mg -MINMATCH=0.1 -LDOFCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10701200 @CNC 1.1 470 @runat.25022005.102832.10300 -NCPU=6 -ICPU=3
-NO WMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FEAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq.16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2284	100.0	1311	6	ABL51513
2	2284	100.0	1311	6	ABL51513 Methylo
3	2284	100.0	1311	6	ABL51513 Methylo
4	1444.5	63.2	8530	8	ABL51513 High grow
5	1444.5	63.2	8530	8	ABL51513 High grow

6	659	28.9	552	13	ADS62361	Ad62361 Bacterial
7	306.5	13.4	1050	13	ADS56121	Ad56121 Bacterial
8	281.5	12.3	1458	3	ACA46084	ACA46084 Arabidops
9	270	11.8	1347	8	ACA23197	ACA23197 Prokaryot
10	270	11.8	1778	2	ACA85983	ACA85983 Flavaria
11	270	11.8	11000	2	MAX20248_01	Continuation (2 of
12	264.5	11.6	1624	2	AAQ85984	AAQ85984 Oryza sat
13	263.5	11.5	1256	6	ADA70655	Ada70655 Rice gene
14	263	11.5	1558	2	AAQ85986	AAQ85986 Raphanus
15	262	11.5	960	13	ADS48162	ADS48162 Bacterial
16	262	11.5	1978	2	AAQ85982	AAQ85982 Solanum t
17	261.5	11.4	1490	8	ADA70675	Ada70675 Rice gene
18	261.5	11.4	1491	10	ADC07917	ADC07917 Rice DNA
19	261	11.4	1035	13	ADT44510	Adt44510 Bacterial
20	258.5	11.3	1769	3	ACA44555	ACA44555 Zea mays
21	257.5	11.3	1503	8	ADA68295	Ada68295 Arabidops
22	257	11.3	1599	8	ADA70610	Ada70610 Rice gene
23	254.5	11.1	1550	8	ADA70520	Ada70520 Rice gene
24	251	11.0	1810	3	ACA44927	ACA44927 Arabidops
25	249.5	10.9	956	13	ADS57682	Ad57682 Bacterial
26	248	10.9	1073	13	ADS57682	Ad57682 Bacterial
27	246.5	10.8	2964	13	ADT47247	Adt47247 Bacterial
28	240.5	10.5	960	8	ACA48934	ACA48934 Prokaryot
29	240.5	10.5	963	4	AA556050	AA556050 Salmonell
30	240.5	10.5	963	8	ACA50979	ACA50979 Prokaryot
31	240.5	10.5	963	8	ACA51971	ACA51971 Prokaryot
32	240	10.5	1053	3	ACA42127	ACA42127 Arabidops
33	238.5	10.4	963	4	AA552688	AA552688 E. coli D
34	238.5	10.4	963	8	ACA322717	ACA322717 Prokaryot
35	238.5	10.4	963	13	ADT48887	Adt48887 Bacterial
36	234.5	10.3	984	8	ACA54036	ACA54036 Prokaryot
37	234	10.2	1530	3	ACA48983	ACA48983 Arabidops
38	232	10.2	1386	4	ACA52344	ACA52344 Prokaryot
39	232	10.2	37337	4	AA559518	AA559518 Propionib
40	232	10.2	37337	8	ACF64447	ACF64447 Propionib
41	230.5	10.1	1164	12	ADJ40285	Adj40285 Plant cDN
42	229.5	10.0	885	13	AA556868	AA556868 Bacterial
43	229.5	10.0	960	8	ACA32326	ACA32326 Prokaryot
44	228	10.0	7972	2	AAV74623	AAV74623 Staphyloc
45	227.5	10.0	1349	13	ADS56889	Ad56889 Bacterial

ALIGNMENTS

RESULT 1	ABL51513	standard; DNA; 1311 BP.
ID	ABL51513	
XX	ABL51513;	
AC		
XX		
DT	01-JUL-2002 (first entry)	
XX		
DE	Methylomonas pyrophosphate dependent phosphofructokinase DNA SEQ ID:15.	
XX		
KW	Methylomonas; methanotropic; carbon metabolism; carbon flux pathway;	
KW	transaldolase; fructose biphosphate aldolase; KHG/KDPG aldolase; food;	
KW	pyrophosphate dependent phosphofructokinase; phosphoglucosamine; feed;	
KW	glucose 6 phosphate isomerase; 6-phosphoglucosamine dehydratase; methane;	
KW	glucose 6 phosphate 1 dehydrogenase; enzyme; methanotroph; methanol;	
KW	single carbon substrate; single cell protein; polysaccharide; thickener;	
KW	isoprenoid; carotenoid pigment; gene; ds.	
XX		
OS	Methylomonas sp.	
XX		
PH		
FT	Key	Location/Qualifiers
FT	CDS	1..636
FT		/*tag= a
FT		/partial
FT		/product= "pyrophosphate dependent phosphofructokinase"
XX		/note= "no start or stop codons given"
XX		
PN	MO200220796-A2.	

PD 14-MAR-2002.
 XX 28-AUG-2001; 2001WO-US026730.
 PF 01-SEP-2000; 2000US-0229906P.
 PR (DUPO) DU PONT DE NEMOURS & CO E. I.
 PA Koffas M, Odom JM, Norton KC, Ye RW,
 PI WPI; 2002-362250/39.
 DR P-PSDB; ABB09168.
 XX New polynucleotide encoding a Methylenomonas sp. carbon flux enzyme useful
 PT for altering carbon flow through methanotrophic bacteria, utilized for
 PT production of single cell protein and commercially valuable
 PT polysaccharides.
 PS Claim 2; Page 66; 73pp; English.
 XX The present invention describes isolated polynucleotides (I) encoding a
 CC Methylenomonas sp. carbon flux enzyme, consisting of: transaldolase;
 CC fructose biphosphate aldolase; KMG/KDPG aldolase; phosphoglucosylase;
 CC pyrophosphate dependent phosphofructokinase; glucose 6 phosphate
 CC isomerase; 6-phosphogluconate dehydratase; and glucose 6 phosphate 1
 CC dehydrogenase enzymes. (i) overexpression is useful for altering carbon
 CC flow through a methanotrophic bacteria, where the bacteria grows on a C1
 CC carbon substrate of methane and methanol, and comprises a functional
 CC Embden-Meyerhof carbon pathway comprising a gene encoding a pyrophosphate
 CC dependent phosphofructokinase enzyme, where the bacteria is preferably
 CC Methylenomonas 16a American Type Culture Collection (ATCC) PTA 2402, and
 CC where the carbon flux gene is: (i) over-expressed on a multicopy plasmid;
 CC (ii) operably linked to an inducible or regulated promoter; (iii)
 CC expressed in an antisense orientation; or (iv) disrupted by insertion of
 CC foreign DNA into the coding region. The manipulated methanotrophs are
 CC useful for the energy efficient conversion of single carbon substrates
 CC such as methane and methanol to commercially useful products in the food,
 CC feed and materials industries, and preferably for the production of
 CC single cell protein, and for producing polysaccharides, used as
 CC thickeners in food and non-food industries, and isoprenoid compounds and
 CC carotenoid pigments of various carbon lengths. The present sequence
 CC encodes a Methylenomonas pyrophosphate dependent phosphofructokinase from
 CC the present invention
 XX
 SQ Sequence 1311 BP; 327 A; 320 C; 366 G; 298 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,78e-229 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-10-701-200-6 (1-437) x ABL51513 (1-1311)
 QY 1 AapvAlvalThrTrProTyRHisLeuThraAlaaspIleargPheCyshiStrPhephe 20
 DB 1 GATGGTGCACATGCGCCATCATCACTTAACGGCTGATTCGATTGTGCTGTTT 60
 QY 21 LeuanpPheasnPheTyRThrLeuMetAenlyPrProLybValAlaIleLeuthraAla 40
 DB 61 CTTAACCTTTAACTTCTACACGCTCATGAACAAACCTAAAGTTGCAATCTGACAGCA 120
 QY 41 GlyGlyLeuAlaProCyLeuAsnSerAlaIleGlySerLeuIleGluArgTrpThrglu 60
 DB 121 GCGGCGCTTGCGCCCTTGTTGATTCCGAAATCGGAGTTTATCGAACGTTATACCGAA 180
 QY 61 IleappProSerIleGluIleIleCyRtyrArgGlyGlyTyRlyGlyLeuLeuLeuGly 80
 DB 181 ATCGATCCCTAGCATGAATCATTTGCTATCGCGGGGTTATTAAGGCGCTGTTGCTGGGC 240
 QY 81 AapSerTyRProValThrAlaGluValArgLybValIleGlyValLeuGlnArgPheGly 100

DB 241 GATTCTTATCCAGTACGCGCAAGTCGTTAAAGCGGGGCTTCTGCAACGTTTGGC 300
 QY 101 GlySerValIleGlyAsnSerArgValIlybLeuThraAsnValIyAspCyValIySarg 120
 DB 301 GGTTCGTGATCGGCAACAGCGCGCTCAAAATTGACCAATGTCAAAAGATCGCTGAAAGC 360
 QY 121 GlyLeuValIlybGlyGlyGluAspProGlnIySValAlaIAspGlnLeuValIySap 140
 DB 361 GGTTCGTCAAAAGAGGCTGAAGATCCGCAAAAGTCGGCGCTGATCATTTGGTTAAGAT 420
 QY 141 GlyValAspIleLeuHisThrIleGlyGlyIyAspAspThrAsnThraAlaIAspLeu 160
 DB 421 GGTTCGATTTATTCGACACCATCGCGCGGATGATACCATACGCGACGACGGAATTTG 480
 QY 161 AlaIAspLeuAlaArgAsnAsnTyRtyRLeuThraValIleGlyLeuProLySThrVal 180
 DB 481 GCAGCATTCCTGCGCGCAAAATTAATTAAGGATCGGATCGCTGTTACTTAAACCGTC 540
 QY 181 AapAsnAapValPheProIleIyGlnSerLeuGlyAlaATPThrAlaIAspGlnGly 200
 DB 541 GATTAACGACGTAATTCGATCAACGCAATCAGTGGCTTGACCTGCGCGCAAGAGC 600
 QY 201 AlaArgTyRProMetAsnValValIAspIySAsnAlaAsnProArgMetLeuIleVal 220
 DB 601 GCGCTTATTTCAATGAACTGTGTGCGCAAAACCAACGCAACGCACTGTGATCGTA 660
 QY 221 HisGluValMetGlyArgAsnCySgIyRProLeuThraAlaIAspGlnIySArg 240
 DB 661 CACGAAGATGATGGCGCGTAATCGCGGCTGACCGCTGCAACCGCGCAAGAAATATGCG 720
 QY 241 LybLeuLeuAapArgAlaGluTrpLeuProGlnIySLeuThraArgGlySerTyRglu 260
 DB 721 AAATTAATCGACCGCGCGAGGTGGTGGCAATGGTGTGACTCGTGAATCTTAATGAA 780
 QY 261 ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaIySArgLeu 280
 DB 781 GTGACGCGGATTCGTTCGCGAAATGCGCATCGCTGAAAGCGCAAGCGCGCTG 840
 QY 281 ArgGluValMetAspLybValIAspCySValAsnIlePheValSerGlyGlyValIyVal 300
 DB 841 CCGCAAGATGATGACAAAGTGGATTCGCTCAACATCTTCGTTCCGAAGTCCGCGCGTC 900
 QY 301 GluAlaIleValAlaGluMetGlnAlaIySgIyGlnIyValProArgAspAlaPheGly 320
 DB 901 GAAGCTATCGTCGCGAAATGACAGGCCAAAGCGCAAGTGCCTCGCATCTGTCGCG 960
 QY 321 HisIleLybLeuAspAlaValAsnProGlyLySTrpPheGlyGlyGlnPheAlaGlnMet 340
 DB 961 CACATCAAACTGATGCGGTCAACCTGTGTAAATGTTTCGGGACCAATTCGCGCAGATG 1020
 QY 341 IleGlyAlaGlyLybThrLeuValGlnIySserGlyTyRPhaAlaArgAlaSerAlaSer 360
 DB 1021 ATAGCGCGGAAAAAACCTCGGTACAAABAATCGGGAATCTTCGCGCGCTTCCTGCTTC 1080
 QY 361 AsnValAapAspMetArgLeuIleIySserCySAlaAspLeuAlaValIyGlySAlaPhe 380
 DB 1081 AACGTTGACGACATGCGTTTATCAAAATCGTGGCGGCTTCGCGGTGAGAGCGCGCTTC 1140
 QY 381 ArgArgGlySerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle 400
 DB 1141 CGCGCGAGTCTGGGTATCGGTCAACAGCAAGCAACGCGACAGTGTGGTGCAGATC 1200
 QY 401 GluPheProArgIleIySgIyGlyLySProPheAsnIleAspThrAspTrpPheAsnSer 420
 DB 1201 GAGTTTCGCGCATCAAGCGCGCAACGCTTCATATCGAACCGATCGTTCAATAC 1260
 QY 421 MetLeuSerGluIleGlyGlnProLySgIyGlyLySValGluValSerHis 437
 DB 1261 ATGTTGACGAAATCGCGCACGCTTAAAGCGGTAAAGTCAAGTCAAGCAC 1311
 RESULT 2
 AAD35497

ID AAD35497 standard; DNA; 1311 BP.
 AC AAD35497;
 XX
 XX
 DT 07-AUG-2003 (revised)
 DT 25-JUL-2002 (first entry)
 XX
 DE Methylomonas 16a sp. phosphofructokinase pyrophosphate dependent gene.
 XX
 XX Carotenoid; isopentenyl pyrophosphate; antheraxanthin; asaxanthin; diet;
 KM anti-oxidant; steroid; flavour; fragrance; electro-optic application;
 KM aquaculture; enzyme; phosphofructokinase pyrophosphate; gene; ds.
 XX
 OS Methylomonas sp.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1311
 FT /tag a
 FT /product= "Phosphofructokinase pyrophosphate"
 FT /note= "CDS does not include start and stop codon"
 FT /partial
 XX
 PN MO200218617-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 04-SEP-2001; 2001WO-US027420.
 XX
 PR 01-SEP-2000; 2000US-0229858P.
 PR 01-SEP-2000; 2000US-0229858P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PI Brzostowicz PC, Cheng Q, DiCosimo DJ, Koffas M, Miller ES;
 PI Odom JM, Picataggio SK, Rouviere PB;
 DR WPI; 2002-351711/38.
 DR P-PSDB; AAE22299.
 XX
 XX Producing carotenoid compounds e.g. antheraxanthin and asaxanthin, by
 PT using microorganisms having a nucleic acid molecule encoding enzymes in
 PT the carotenoid biosynthetic pathway and which metabolize single carbon
 PT substrates.
 XX
 XX
 PS Example 2; Page 103-104; 156pp; English.
 XX
 CC The invention relates to a method for producing carotenoid compounds. The
 CC method comprises a transformed metabolising host cell, comprising
 CC suitable levels of isopentenyl pyrophosphate and a nucleic acid molecule
 CC encoding an enzyme in the carotenoid biosynthetic pathway, under the
 CC control of regulatory sequences, and contacting the host cell with carbon
 CC substrate to produce a carotenoid compound. The method is useful for
 CC producing carotenoid compounds such as antheraxanthin and asaxanthin, by
 CC using microorganisms having a nucleic acid molecule encoding enzymes in
 CC the carotenoid biosynthetic pathway and which metabolize single carbon
 CC substrates. The carotenoids have potent anti-oxidant properties useful in
 CC diet, and aquaculture elements. The carotenoids are also useful as
 CC intermediates in the syntheses of steroid flavours and fragrances and
 CC compounds for potential electro-optic applications. The present sequence
 CC is Methylomonas 16a sp. phosphofructokinase pyrophosphate dependent
 CC enzyme gene used in the invention. (Updated on 07-AUG-2003 to correct OS
 CC field.)
 CC
 SO Sequence 1311 BP; 327 A; 320 C; 366 G; 298 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,78e-229
 Score: 2284.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 6

Length: 1311

Matches: 437

Conservative: 0

Mismatch: 0

Indels: 0

Gaps: 0

US-10-701-200-6 (1-437) x AAD35497 (1-1311)

QY	1	AspValValThrTrpProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhe	20
DB	1	GATGTGTCACATGGCCCTATCATCTTAACGGCTGATTCGATTTTCTCATTTGTTT	60
QY	21	LeuAsnPheAsnPheTyrThrLeuMetAsnLysProLysIleValAlaIleLeuThrAla	40
DB	61	CTTAACCTTAACCTTACACAGCTCATGAACAAACCTAAAGTTCCAAATACGACAGCA	120
QY	41	GlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGlu	60
DB	121	GGCGGCTTGCGCGCTTGTGTAATTCGCAATCGTAGTTGATTCGAACCTTATACCGAA	180
QY	61	IleAspProSerIleGluIleIleCysTyrArgGlyGlyTyrArgGlyLeuLeuGly	80
DB	181	ATCGATCCTAGCATTAATAATCATTTCTATCGCGGCTTATTAAGCCCTGTTCGCGC	240
QY	81	AspSerTyrProValThrAlaGluValArgLysLysAlaGlyValLeuGlnArgPheGly	100
DB	241	GATTTCTTATCCAGTAAAGCCGCAAGTCGTAATAAAGCGGTGTCGCAACGTTTGGC	300
QY	101	GlySerValIleGlyLysSerArgValLysLeuThrAsnValLysAspCysValLysArg	120
DB	301	GGTTCTGTGATCGCAACAGCCGCTCAATATGACCAATGTCAAGACTGCGTGAACGC	360
QY	121	GlyLeuValIleGlyGlyGlyLeuAspProGlnLysValAlaAlaAspGlnLeuValLysAsp	140
DB	361	GGTTTGTGTAAGAGGTGAAAGTCCGCCAATAAGTCCGGCTGATCAATTTGGTTAAGAT	420
QY	141	GlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAlaAlaAspLeu	160
DB	421	GGTGTGATATTCTGCACACCATCGCGGATGATACCAATACGCGACAGCGGATTTG	480
QY	161	AlaAlaPheLeuAlaArgAsnSerTyrGlyLeuThrValIleGlyLeuProLysThrVal	180
DB	481	GCAGCATCTCTGCGCGCAATAATTAACGATCACTGCTTATCTTAACCTTAACCGTC	540
QY	181	AspAsnAspValPheProIleLysGlnSerLeuGlyAlaTrpThrAlaAlaGlnGly	200
DB	541	GATTAACGATATTTCGATCAACGATCACTAGTGTCTTGATCTCCGCGCAAGGC	600
QY	201	AlaArgTyrPheMetAsnValValAlaGluAsnAsnAlaAsnProArgMetLeuIleVal	220
DB	601	GCGGTTATTCTATGAACGTCGGTGGCGCAAAACACGCCAACCCACGATGATCTGTA	660
QY	221	HisGluValMetGlyArgAsnCysGlyTrpLeuThrAlaAlaThrAlaGlnGlyTyrArg	240
DB	661	CACGAAGTATGGCGCGTAATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720
QY	241	LysLeuLeuAspArgAlaGluTrpLeuProGlnLeuGlyLeuThrArgLysSerTyrGlu	260
DB	721	AAATTAATGACCGTCCGAGTGTGTCGGAATGGGTTTGAATCTGTGAATCTTATGAA	780
QY	261	ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluAlaLysArgLeu	280
DB	781	GTCACAGCGGTATTTCTTCGGAATGGGATGATGATGATGATGATGATGATGATGATG	840
QY	281	ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluValIleGlyVal	300
DB	841	CGGGAAGTATGACCAAGTCGATTCGTCACATCTTCTGTTCCAAAGTTCGCGCTC	900
QY	301	GluAlaIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly	320
DB	901	GAAGCTATCGTCGCGAATATGACGCCAAGAGCCAGAAAGTCCGCGATGCTTCGCGC	960
QY	321	HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGlyGlnPheAlaGlnMet	340
DB	961	CACATCAAACTGATCGCGTCAACCTGTGAATATGTTGGCGACGATTCGCGCAGATG	1020
QY	341	IleGlyAlaGluLysThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSer	360
DB	1021	ATAGGCGCGGAAAAAACCCTGTGTAACAAATATCGGATACCTTCCCGCTGCTTCC	1080

QY 361 AaVa1aAaPAPMeCaRgLeu1leYsSerCyAa1aAaPLeu1aVa1G1uCySa1aPhe 380
DB 1081 AACGTTGACGACATGCGTTGATCAAAATGTCGCGCACTTGCGGTGACAGCGCGTTC 1140
QY 381 AaGaRgG1uSeRg1yVall1eG1yH1aSPG1uAaPaRg1yAaPaVall1eUaRg1a1le 400
DB 1141 CCCCCGAGTCGCTGATCGCTACGACGAAGACGCAACGCTGTCGCTGATC 1200
QY 401 G1uPheProA1g1eYsG1yLysProPheAa1leAaPThraSPThPheAaSeR 420
DB 1201 GAGTTTCGCGCATCAAGCGCGCAACCGTTCAATATCGACACGACGTTCAATAC 1260
QY 421 MeTLeuSeRg1u1leG1yG1uPProLySg1yG1yLysVa1G1uVa1SeRHiS 437
DB 1261 ATGTTGAGCAATCGCGCACGCTTAAAGCGGTAAAGTGAAGTCAGCCAC 1311

RESULT 3
ABK83232
ID ABK83232 standard, DNA, 1311 BP.
XX
AC ABK83232;
DT 27-AUG-2002 (first entry)
XX
DE High growth methanotrophic bacterial strain DNA #3.
XX
KW High growth methanotrophic bacterial strain; C1 carbon substrate; gene;
KW methane; methanol; Embden-Meyerhof carbon flux pathway; 16S rRNA; ds;
KW pyrophosphate dependent phosphofructokinase; nitrogen-containing compound;
KW ammonia; nitrate; nitrite; nitrogen; pigment; oxygen; landfill;
KW methane-containing environment; waste water treatment system; isoprenoid;
KW nitrous oxide; terpenoid; animal feed; carotenoid; exopolysaccharide.
XX
OS Methylocomas sp.
XX
PN MO200220728-A2.
XX
PD 14-MAR-2002.
XX
PF 28-AUG-2001; 2001WO-US026827.
XX
PR 01-SEP-2000; 2000US-0229858P.
XX
PI (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PA Koffas M, Odom JM, Schenzle A;
XX
PI MPI: 2002-454358/48.
XX
DR P-PSDB; ABG61553.
XX
PT New high growth methanotrophic bacterial strain, useful for producing
PT single cell proteins, grows on a C1 carbon substrate, and comprises a
PT functional gene encoding in Embden-Meyerhof carbon pathway.
XX
PS Claim 29; Page 85; 157pp; English.
XX
XX The invention relates to a high growth methanotrophic bacterial strain,
CC which grows on a C1 carbon substrate e.g. methane and methanol, and
CC comprises a functional Embden-Meyerhof carbon flux pathway comprising a
CC gene coding a pyrophosphate dependent phosphofructokinase enzyme or a 16S
CC rRNA. The bacterial strain is useful for the production of single cell
CC protein and for the biotransformation of a nitrogen-containing compound,
CC e.g. ammonia, nitrate, nitrite or nitrogen. It is also useful for the
CC production of a feed product comprising a protein, carbohydrates and a
CC pigment and for reducing oxygen demand, for removing nitrates and
CC nitrates in methane-containing environments such as landfill, waste
CC water treatment systems or anywhere that methane, oxygen and nitrates are
CC present. The bacterial strain of the invention can be used as a
CC denitrifying agent for the conversion of nitrate or nitrite to nitrous
CC oxide with methane or methanol as a carbon source. It is also used in the
CC production of biomass including proteins, carbohydrates and a wide
CC variety of pigments (particularly for isoprenoid pigments for the purpose

CC of generating animal feeds), in production of terpenoid and carotenoid
CC compounds, useful as pigments and as monomers in polymeric materials and
CC in production of exopolysaccharides at high levels. Sequences ABK83230-
CC ABK83270 represent high growth methanotrophic bacterial strain DNA of the
CC invention
XX
SQ Sequence 1311 BP, 327 A, 320 C, 366 G, 298 T; 0 U, 0 Other;
Alignment Scores:
Pred. No.: 1,786-229 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-701-200-6 (1-437) x ABK83232 (1-1311)
QY 1 AaPVal1a1Thr1TrpProTyrThiA1eUhr1aAaP1leA1rgPheCySeH1a1TrpPhePhe 20
DB 1 GATGTGTCACATGAGCCCTATCATCTTAAAGCTGATATTCGATTTGTCATTGTTT 60
QY 21 LeuAaPheAaPheTyrThrLeuMeCaRg1yLysProLyS1yVa1a1a1leUhr1a1a 40
DB 61 CTTAaCTTAACTTCTTACACGCTCATGAACAACTTAAAGTTCGAATACTGACAGCA 120
QY 41 G1yG1yLeuA1aProCyA1eUaAaSeR1a1leG1ySeR1eU1eG1uA1gTyrThrg1u 60
DB 121 GCGGCTTGCGCGCTGTTGTAATCCCAATCGGTAGTTGATGAAGTATACCGA 180
QY 61 l1eAaPProSeR1leG1u1le1eCySeR1yA1rg1yG1yLys1yUeUeUeUg1y 80
DB 181 ATCGATCTTACATGAATGAATCATTTGCTATCGCGGCTTAAAGGCTGTGCTGCGC 240
QY 81 AaPSeR1yProVa1Thra1aG1uVa1A1rg1yLys1a1aG1yVa1leUg1na1rgPheG1y 100
DB 241 GATTCTTATCAAGTAAACGCGCAAGTACGCTTAAAGGCGGCTTCTGCAAGCTTTGGC 300
QY 101 G1ySeR1a1leG1yAaSeR1yA1rg1yLys1eUhr1aAaP1yAaP1yVa1rg 120
DB 301 GGTTCGTGATCGGCAACAGCGCGCTCAAAATGACCAATGTCAAGACTGCGTAAGCC 360
QY 121 G1yLeuVa1ySg1uG1yG1uAaPProG1u1yS1a1a1aAaP1leUeU1yAaP 140
DB 361 GGTTCGTCAAGAGAGGTGAAGATCCGAAAAAGTCGCGCTGATCAATGTGTTAAGAT 420
QY 141 G1yVa1aP1leUeU1eThr1leG1yG1yAaPAPThraSPThra1a1a1aAaP1leU 160
DB 421 GGTTCGATATTTCTGACACCATTCGCGCGCATGATACCAATACGACGACGATTTG 480
QY 161 A1a1aPheUeU1aA1rgAaAaNTyrg1yLeuThra1a1leG1yUeUProLyS1yThra1 180
DB 481 GGAGCATTCCTGCGCAGAAATATTAACGACGACGCTATGTTTACCTTAAACCGTC 540
QY 181 AaPaAaPVal1PhePro1leYsG1uSeR1eUg1yA1a1TrpThra1a1aG1uG1ng1y 200
DB 541 GATTAACGACGATTTCCGATCAAGCAATCATAGGTGCTTGACCTGCCGCAAGCAAGGC 600
QY 201 A1aA1gTyrPheMeCaRg1yVa1a1aG1uAaAaNTyrg1yAaPProLyS1yUeU1eUa1 220
DB 601 GCGCTTATTTTCATTAACGCTGCGCGCAAAACCAACGCAACGCAATGTCGATCGTA 660
QY 221 H1aG1uA1MeG1yA1rgAaNTyrg1yTrpLeuThra1a1a1eThra1aG1ng1yY1rg 240
DB 661 CACGAAGGATGAGCGCGTAACCTGCGGTGACCGCTGCAACGCGCAGGAATATCCG 720
QY 241 Lys1eUeUeUaP1yA1aG1uTrpLeuProG1uLeUeUg1yLeUThra1rg1uSeR1yrg1u 260
DB 721 AAATTAAGGACCGCGGAGAGGTGCGGGAATGGGTTTGAATCGTGAATCTTAATGA 780
QY 261 VAlH1a1a1a1PheVa1ProG1uMeCa1a1eAaPLeUg1uA1aG1uA1a1yA1rg1u 280
DB 781 GTGACGCGGTATTCGTTCCGAAATGCGGATCGACTGAAAGCCGAAGCCAGCCCTG 840

QY 281 ATGGTAAAGTGTGACAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 300
DB 841 CGGGAAGTGTGACAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 900
QY 301 GUAATTTTAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 320
DB 901 GAAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 960
QY 321 HATTTTAAAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 340
DB 961 GATTTTAAAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 1020
QY 341 TTTTAAAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 360
DB 1021 ATGAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 1080
QY 361 AATTTTAAAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 380
DB 1081 AATTTTAAAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 1140
QY 381 ATGAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 400
DB 1141 CGCGGAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 1200
QY 401 GATTTTAAAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 420
DB 1201 GATTTTAAAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 1260
QY 421 MetLeuSerGluIleGlyGlnProlyGlyGlyValGluValSerHis 437
DB 1261 ATGTTGAGCGAATTCGCGCCAGCTTAAAGCGGTAAAGTGCAGAGTCAGCCAC 1311

RESULT 4

AAS59528
ID AAS59528 standard; DNA; 8530 BP.

AC AAS59528;

DT 13-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein encoding DNA #23.

KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant; de.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-019047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (COR1-) CORIXA CORP.

PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'mabonneuve J, Zhang Y, Jen S, Carter D;

DR WPI; 2001-616774/71.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT creating acne vulgaris.

XX Claim 1; SEQ ID NO 23; 10699P; English.

CC Sequences AAS59506-AAS59804 represent DNA molecules encoding
CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
CC associated DNA sequences are used in the treatment, prevention and
CC diagnosis of medical conditions caused by P. acnes. The disorders include
CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in
CC infections of bone, joints and the central nervous system, however it is
CC particularly involved in the inflammatory lesions associated with acne
CC vulgaris. A method for detecting the presence or absence of P. acnes in a
CC patient comprises contacting a sample with a binding agent that binds to
CC the proteins of the invention and determining the amount of bound protein
CC in the sample. The polypeptides may be used as antigens in the production
CC of antibodies specific for P. acnes proteins. These antibodies can be
CC used to downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
CC polypeptides shown in AUA4543-AUA45926 and AUA67500. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SO Sequence 8530 BP; 1921 A; 2633 C; 2423 G; 1551 T; 0 U; 2 Other;

Alignment Scores:

	Pred. No.:	Length:	Matches:
Score:	2,72e-140	8530	274
Percent Similarity:	144.50		274
Best Local Similarity:	80.59%		54
Query Match:	67.32%		74
DB:	4	Indels:	5
		Gaps:	1

US-10-701-200-6 (1-437) x AAS59528 (1-8530)

QY 33 LysIleValAlaIleLeuThrAlaGlyIleValAlaProCyteLeuAsnSerAlaIleGly 52
DB 3962 AAGAAAGTGTGACAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTCAGT 4021
QY 53 SerLeuIleGluArgThrGluIleAspProSerIleGluIleCysTyraArgly 72
DB 4022 GATTTTAAAGTGTGATTCGTTCCGAAAGTGTGATTCGTTCCGAAAGTGTGCGCGTC 4081
QY 73 GlyTyraGlyLeuLeuLeuGlyAspSerTyraProValThrAlaGluValArglyLys 92
DB 4082 GGTAAAGAGGCGCTGTTCAAGGAGCTTCGAGAGTCAACGACCGGTGCAAGAAT 4141
QY 93 AlaGlyValLeuGlnArgPheGlyIleSerValIleGlyAsnSerGlyValLysLeuThr 112
DB 4142 GCCGAGATCCTCAAGCGTTTGGTGTGATCCCGATTGGCAACTCGCGGTCAAGCTGACC 4201
QY 113 AsnValIleAspCysValIleAspArgGlyLeuValIleGluGlyLysAspProGlnLysVal 132
DB 4202 AACGCCCGGACCTGTAAGCGTGTGATTCGTTCCGAAAGTGTGCGCGTCAGT 4261
QY 133 AlaAlaAspGlnLeuValIleAspArgGlyValAspIleLeuHisThrIleGlyIleAspAsp 152
DB 4262 GCGCGGATCGTGTGTTCCGACGCGGTGATTTTGCACACCATCGGTGACGAC 4321
QY 153 ThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTyraGlyLeuThr 172
DB 4322 ACTAATACCAACCGCCCGCATCGCGGTGATTCGTTCCGAAAGTGTGCGCGTCAC 4381
QY 173 ValIleGlyLeuProIleThrValIleAspAspValPheProIleGlyGlnSerLeuGly 192
DB 4382 GTTCGTGCGCTGCGAAGACCATCGATACGACGTGCGCATTCGTTCCGAAAGTGTGCGCGTC 4441
QY 193 AlaIleThrAlaAlaGluGlnGlyAlaArgTyraPheMetAsnValAlaGluAsnAsn 212
DB 4442 GCTTGAACGCGTGGAAACAGGTTCCCGCTTGTCTGCAACCATCGTTGCGAGCAAC 4501
QY 213 AlaAsnProArgMetLeuIleValIleGluValMetGlyArgAsnGlyArgTyraPheLeuThr 232
DB 4502 TCCGATCAACGATGCTCATCGTCAAGAGTCAATGAGCGGTAACTGCGGTGCTCACC 4561

QY	233	AlaAaTtRrAlaGInGluTyrArgGlyLeuLeuAspAgaGAlaGTrpLeuProGluLeu	252
Db	4562	GCcCtTACcCGcCGcCAATtACCGGAAtGGcCTCATtCTCAACAGtGGcCTtACtGAGtC	4622
QY	253	GlyLeuThraArgJuseTyrGluValHAlaValaPheValProGluMetAlaIleAsp	272
Db	4622	GGGcGTtCCAAAGAGGcCTGGAGtGTCCACcCGGATtATGTtCCCGAGGcTCACATtCGAC	4681
QY	273	LeuGluAlaGluAlaLysArgLeuArgGluValMetAspLysValAspCysValAsnIle	292
Db	4682	CTTAAGGcCGAAGcCGACGcGcCTCAACAGGtCATGAGtGGAAGtTGGCAAGcCTCATtC	4741
QY	293	PheValSerGluGluAlaGlyValaGluValaIleValaIaGluMetGlnAlaLysGlyGln	312
Db	4742	TTcCTGTtCCGAGGGGGcGTGGcCTGGAGcCATtTtTAGAGAGtTGAGAAAAGAcGTCTAG	4801
QY	313	GluValaProArgAspAlaPheGlyHAlaLysLeuLeuAspAlaValaAsnProGlyLysTrp	332
Db	4802	GAAcTtCCCGcCGcCGATtCCcCTTtTGtTCAcCTCAAGcCTCGACAAGtGTAAcCCGGtGTcCTTGG	4861
QY	333	PheGlyGluGlnPheAlaGlnMetIleGlyValaGluTyrAsnThraLeuValGlnLysSerGly	352
Db	4862	TTGGcCAAAcAGtTtCGcCGACAGcTcGtGcCGAAAAGtCATGtGTCAAGtTCCGGcC	4921
QY	353	TyrPheAlaArgAlaSerAlaSerAsnValaAspAspMetArgLeuIleLysSerCysAla	372
Db	4922	TACtTtTtCCcGTtGGcAGcCTCGAACAAGcCTGACtTGAGAGtCATtGGcCGcCTGACT	4981
QY	373	AspLeuAlaValaGluCysAlaPheArgArgGluSerGlyValaIleGlyHAlaAspGluAsp	392
Db	4982	GACcTtCGcGTtCGATtTGcCGcCTGGcCGGTAAcTGGcCGtTATtGGcCAGAGcCAGAG	5041
QY	393	AsnGlyAsnValaLeuArgAlaIleGluPheProArgIleLysGlyGlyAspProPheAsn	412
Db	5042	AACGGcCAcAGcCTTGACAGACAATGcCTTCATtCGAGtCAAGGGcCGCAAGcCTTTCAGC	5101
QY	413	IleAspThraAspTrpPheAsnSerMetLeuSerGluIleGlyGln-----	427
Db	5102	ACCAcTcAGcCGGTGtTCAcCGcCATGcCTCAcCGAGATcGGGcAGcGcCTGAcAGcTATGG	5161
QY	428	ProLysGlyGlyLysValaGlu 434	
Db	5162	CCCAACAGtGGcCCTTCGAA 5182	
RESULT 5			
ACF64457	ACF64457 standard; DNA; 8530 BP.		
XX	ACF64457,		
AC			
XX			
DT	17-OCT-2003 (first entry)		
XX			
DE	Propionibacterium acnes DNA contig sequence #23.		
XX			
KW	Acne vulgaris; anti-seborrheic; dermatological; antibacterial;		
KW	immunostimulant; immune response; vaccine; ds.		
XX			
OS	Propionibacterium acnes.		
XX			
PN	WO2003033515-A1.		
PD			
XX	24-APR-2003.		
PD			
XX			
PF	11-OCT-2002; 2002WO-US032727.		
XX			
PR	15-OCT-2001; 2001US-00978825.		
XX			
PA	(CORI-) CORIXA CORP.		
XX			
PI	Mitchem JJ, Skelley YAW, Persing DH, Bhatia A, Maisonneuve JJ,		
PI	Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;		
PI	Batch B, Valliere-Douglas J;		

XX	WP1; 2003-381789/36.
XX	New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT	polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
FT	or for stimulating an immune response specific for a P. acnes protein.
XX	Claim 1, SEQ ID NO 23, 1481bp; English.
PS	The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC	encoding a Propionibacterium acnes protein. The invention also relates to
CC	polypeptides encoded by the polynucleotides (ABM55624-ABM64536) and to
CC	immunogenic fragments of P. acnes polypeptides. The invention
CC	additionally encompasses expression vectors and host cells comprising a
CC	polynucleotide of the invention; antibodies against polypeptides of the
CC	invention; fusion proteins comprising a polypeptide of the invention; a
CC	method for stimulating an immune response specific for a P. acnes
CC	polypeptide and an isolated T cell population comprising T cells prepared
CC	via this method; a vaccine composition (comprising P. acnes polypeptides,
CC	polynucleotides, antibodies, fusion proteins, T cell populations, or
CC	antigen-presenting cells that express the polypeptide), a method and kit
CC	for detecting or determining the presence or absence of P. acnes in a
CC	patient; and a method for inhibiting the development of P. acnes in a
CC	patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC	proteins, T cell populations or antigen-presenting cells that express the
CC	polypeptides are useful for diagnosing, preventing or treating acne
CC	vulgaris, or for stimulating an immune response specific for a P. acnes
CC	protein. The polynucleotides can also be used as probes or primers for
CC	nucleic acid hybridisation. The vaccine composition is useful for the
CC	stimulation of an immune response against P. acnes, or for treating acne,
CC	and the kit is useful for performing a diagnostic assay. The present
CC	sequence represents a P. acnes DNA contig which is specifically claimed
CC	in the invention. Note: The sequence data for this patent did not form
CC	part of the printed specification, but was obtained in electronic format
CC	directly from WPO at ftp.wipo.int/pub/published_pct_sequences
CC	
XX	
SQ	Sequence 8530 BP; 1921 A; 2633 C; 2423 G; 1551 T; 0 U; 2 Other:
Alignment Scores:	
Pred. No.:	2,72e-140 Length: 8530
Score:	1444.50 Matches: 274
Percent Similarity:	80.59% Conservative: 54
Best Local Similarity:	67.32% Mismatches: 74
Query Match:	63.24% Indels: 5
DB:	Gaps: 1
US-10-701-200-6 (1-437) x ACF64457 (1-8530)	
OY	33 lyelyelvalalalleleuthrralaglglyleunlabProCysleuanSerAlallegly 52
Db	3962 AAGAAAGTTCGATCCCTCACCGCGGGGGAATTGGCGCGGTCCATTCACCGCATTCAGT 40221
OY	53 SerleutlegluarqgryrThrgluilaaspProserilegullieleCyTYrArGoly 72
Db	4022 GGATTGATCCAGCGCTACACCGAGTGACACCTAGGTCAAGATCATGCGTTACAAGAC 4081
OY	73 GlyTyrlrsglyleuleuleuglyyAspSerryrProvalThrAlaGlValAxlLylys 92
Db	4082 GGCTACAGAGGCGCTGCATCAAGGAGACTTCCTCGAGGTCAACCGACCGTTCGCAAGAT 4141
OY	93 AlaglyValleuElnAgrPhegllyylSerValillellymnserrArgValyleuThr 112
Db	4142 GCCCAGATTCCTCAAGCTTTTGGTGGATGCCCGATTGGCAAATCGCGGGTCAAGCTGACC 4201
OY	113 AsnVallyAspCyceVallylsArsgLYleuVallysluglylguaapProGLIlyeVal 132
Db	4202 AACGCCCGCACTCCGTAAGCGGTGGTGTGCTGCCCAAGCGACGACCGCGCTCAAGTT 4261
OY	133 AlaAlaaspGluLeuVallyAspGLyValAspilleleuthrrilleglylglyAspAp 152
Db	4262 GC CGCGCATGCTGCTGCTGGCGACGGGGTTGACATTTTGCACACCATCGGTGGTAGCAC 4321
OY	153 ThrAnthrAlaAlaAlaAspLeuAlaAlaPheleuAlaargbaenrryGlyleuThr 172

```

DB 4322 ACTAATACACACCGCCCATCTGGCCGCTGAGAAATTAACCGCCCTCAC 4381
QY 173 VALLIAGLLeuProlyThrValAaPaaAaPValPheProIleYbGInserLeuGly 192
DB 4382 GTCGTGGCTCCCGCAAGACCATCATTAACGACGTCGTCCTGATTCGTCATCGTAA 4441
QY 193 AATPTThraAlaAglInglValAaArgTYrPheMeCAsnValValAaGluAaen 212
DB 4442 GCTTGACAGCGCTCGCAACAGGTTCCCGCTTGTCTGACAAACATCTGTGGCGACCAAC 4501
QY 213 ALaenProArgMeCLeuIleValAaGluValMeCGLYArgAsnCysGLYTrPLeuThr 232
DB 4502 TCCGATCAACGATGTCATCTGTCACGAAATCATGGCGCTAACTGCGCTGCTCAC 4561
QY 233 ALaAaThraAlaGlnGluTrArgLYLeuAaPArgAlaGluTrPLeuProGluLeu 252
DB 4562 GCCGCTACCGCGCCCAATACCGCAATGAGCTGATCACTCAACAGTGAACCTGAGATC 4621
QY 253 GLYLeuThraArgGluSerTYrGLYValAaAaValPheValProGluMeCAlaAaP 272
DB 4622 GGGCTGTCCAAAGAGCGCTGGAGTGTCCAGCGTATATGTCCCGAGGCTCATCTGAC 4681
QY 273 LeuGluAlaGluAlaLYbArgLeuArgGluValMeCAsPbLYbValAaPcYbValAa 292
DB 4682 CTGAGCGCCGAAGCCCGACCTCAACAGATCATGACAGAAAGTTGCAACGTCATATC 4741
QY 293 PheValSerGluGlyValAglValAaAaValAaGluMeCGLYValAaGluGlyIn 312
DB 4742 TTCCTGTCCGAGGGGCTGCGCTGAGCGCATTAATGAGAGATGGAAGAAAGCGTCA 4801
QY 313 GLUValProArgAaPValAaPheGlyAaIleLYLeuAaPAlaValAaPProGlyLYb 332
DB 4802 GAAGTCCCGCGGATCCCTTGTGTCATGCAAGCTGACAAAGTGAACCGGCTGCTTG 4861
QY 333 PheGlyGluGlnPheAaGlnMeCIIeGLYValAaGluLYbThreValAaGlnLYbSerGly 352
DB 4862 TTCGGCAACAGTTCGCCCAACAGCTCGGTGCGCAAAAGTCTGTCAGAAAGTCCGC 4921
QY 353 TYrPheAlaArgAlaSerAlaSerAaAaValAaPaaPMeCArgLeuIleLYbSerCYbAla 372
DB 4922 TACTTCTCCCGTGGAGCTGCAAGAACTGACGTGAGCTATCGCGCTGCACT 4981
QY 373 AaPLeuAlaValAaGluCYbAaPheArgGluSerGLYValIleGlyAaPGLYbAaP 392
DB 4982 GACTCTGCTGTGATTCGCCCGCGGGAAGACATCGGCTTATCGGCCAGACGAGAG 5041
QY 393 AaNGLYAaAaValLeuArgAlaIleGluPheProArgIleLYbGlyLYbProPheAa 412
DB 5042 AACCGCGACACGTTGACGAACATCGCTTCGATCGGATCAAGGCGGCAAGCTTTCGAC 5101
QY 413 IleAaPThraPTrPheAaAaSerMeCLeuSerGluIleGlyGln----- 427
DB 5102 ACCACTCAACCGCGTTCACCGCATGCTCAGGAGATCGGAGCGGCTGACACTATG 5161
QY 428 ProLYbGLYLYbValGlu 434
DB 5162 CCCAACAGTGGCCCTCCGA 5182

```

RESULT 6
 ADS62361
 ID ADS62361 standard; cDNA, 552 BP.
 XX
 AC ADS62361;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #14348.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; omomais;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;

```

KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomanan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI, 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1, SEQ ID NO 38035; 122pp; English.
XX
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned
CC to provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomanan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 552 BP; 109 A; 181 C; 166 G; 96 T; 0 U; 0 Other;
XX

```

Alignment Scores:

Pred. No.:	Length:	Score:
1.63e-59	552	659.00
Percent Similarity:	Matches:	82.61%
Best Local Similarity:	Mismatches:	67.93%
Query Match:	Indels:	28.85%
DB:	Gaps:	0

US-10-701-200-6 (1-437) x ADS62361 (1-552)

```

QY 33 LYLbValAlaIleuThraAglGlyLeuAlaProCYbLeuAaAaSerAlaIleGly 52
DB 1 CAGAAAGTGCATGTGACCGCGGTGCTCGCGCTGCTCTCTTCCTGCGCGCGGT 60
QY 53 SerLeuIleGluArgTYrThraGluIleAaPProSerIleGluIleCYbTYrArgGly 72

```


Db	61	GGCCCTGATGAGCGCCTATATGCGACATCGCGCCCTGAATGCACATCATCGCTTACCGCTCC	120
Qy	73	GjYrYLygYlYleuDeuLeuGjYasPserTYrProValThzIaGjValaJrYbLYs	92
Db	121	GGTTATCAAGGGGGCTGCTTCTGGGCGAGCGPATCGAGATCACCAAGATATATGGCGAAAG	180
Qy	93	AlaGjYValLeuGlnArGPheGjYgJYserValIlegIYAsnSerArYValYbLeuThr	112
Db	181	GGGCATCTTACTCATGTGTATACGGCGGCTCGCCATGGCGAACACGCGCGTCACGCTGACC	240
Qy	113	AasnValLYasPcyAvValLYsArGjYlYleuValYlYgJYubAPrGJInLYsVal	132
Db	241	AATGCCGCCGATTGCCCGAACCGCGGCTCTGTAAAGAAAGCGACATCCGTTGCCGCTC	300
Qy	133	AlaAlaSPGlnLeuValLYsAsPdjYValAsPjIleuHjSerIlegIYgJYasPasp	152
Db	301	GGCGGTGAACGGCTGGCGGCTGACGGCATACCATTTCTCATATCTATCGGTGGCGAGAC	360
Qy	153	ThrAsnThrAlaAlaAlaAsPleuAlaAlaPheLeuAlaArGAsnAntYrGjYleuThr	172
Db	361	ACCAACACACACCGCGCGGACCTTGGCGCTTACTCGCGCGCCCAACGATATGACCTCAC	420
Qy	173	ValIlegIYleuProLYerThrValAsPAsnAsPValPheProIleLYsgInSerIleuGjY	192
Db	421	GTGCTGGGCTGCGCGAAGACGGTGGATTAATGACGTGCTGCGATCAAGACGTGCTGGC	480
Qy	193	AlATrPThrAlaAlaGJInGjYAlaArGYrPheMeCAsnValYAlaIaGJAsnAsn	212
Db	481	GGCTGGACGGCGCGGAGATCGGTGCTGCTTCTTGACAACTGTCAAGACAGACGAGAGC	540
Qy	213	AlaAsnProArG 216	
Db	541	GCGGACCGAAA 552	
RESULT 7			
ADSS6121			
1D	ADSS6121	standard, cDNA, 1050 BP.	
XX	AC	ADSS6121;	
XX	XX		
DT	02-DEC-2004	(first entry)	
DE		Bacterial polynucleotide #8108.	
XX			
KW		Recombinant DNA construct; transformed plant; improved plant property;	
KW		cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;	
KW		pathogen tolerance; pest tolerance; plant disease resistance;	
KW		cell cycle pathway modification; plant growth regulator;	
KW		homologous recombination; seed oil yield; protein yield; carbohydrate;	
KW		nitrogen; phosphorus; photosynthesis; liginin; galactomanan;	
XX		bacterial polynucleotide; gene; ss.	
OS		Bacteria.	
XX			
PN	US2003233675-A1.		
XX			
PD	18-DEC-2003.		
XX			
PF	20-FEB-2003; 2003US-00369493.		
XX			
PR	21-FEB-2002; 2002US-0360039P.		
XX			
PA	(CAOY/) CAO Y.		
PA	(HINK/) HINKLE G J.		
PA	(SLAT/) SLATER S C.		
PA	(CHEN/) CHEN X.		
XX	(GOLD/) GOLDMAN B S.		
XX			
PI	CAO Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;		
XX			
DR	WPI, 2004-061375/06.		
PT		New recombinant DNA construct comprising a promoter positioned to provide	

PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties
 XX
 PS Claim 1; SEQ ID NO 31795; 122bp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomanan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 1050 BP; 187 A; 374 C; 340 G; 149 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	3,76e-22
Score:	366.50
Percent Similarity:	42.49%
Best Local Similarity:	26.46%
Query Match:	13,42*
DB:	13
Gaps:	14
Length:	1056
Matches:	104
Conservative:	63
Mematches:	145
Indels:	81

US-10-701-200-6 (1-437) x ADS56121 (1-1050)

[illegible]

PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144684P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 23-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145818P.
 PR 27-JUL-1999; 99US-0145819P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 03-AUG-1999; 99US-0146389P.
 PR 04-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 05-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147320P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 09-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 10-AUG-1999; 99US-0147935P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.

PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158025P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158363P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159322P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159844P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 2,47e-19 Length: 1458
 Score: 281.50 Matches: 109
 Percent Similarity: 41.56% Conservative: 61
 Best Local Similarity: 26.65% Mismatch: 166
 Query Match: 12.32% Indels: 73
 DB: 3 Gaps: 16

US-10-701-200-6 (1-437) x MAC46084 (1-1458)

QY 37 l l e u t h r a l a g l y l e u n l a p r o c y s l e u n s e r a l l i e g l y s e r l e u l l e g t u 56
 DB 286 A T A G T T A C A T G G A G G T C T C T G T C C G C T C A T T A C C G T T A G A G A A C T T G T G A C 345
 QY 57 A r g T y r T h r G l u i l l e a s p p r o s e r l l e g l u i l l e i c y s t y r a r g l y g l y t y r y s g l y 76
 DB 346 A G C T T A T C A T A C A T A T T A T G A G T A A G A G A A T T T G G A A T C G A T G G A T T A C A G A G A 405
 QY 77 l e u l e u l e u g l y a s p s e r t y r p r o v a l t h r a l a g l u a l a r g l y a l a g l y v a l l e u 96
 DB 406 T T T T A G C G C A A G A A A T C A T C C C C T T A A C T T A A G T C ----- G T G A A T G A T A T C 456
 QY 97 G l n a r g p h e g l y l y s e r v a l l i e g l y a n s e r a r g v a l l y s l e u t h r a n v a l l y a s p 116
 DB 457 C A T A A C G C A G A G A A C A T A T T G G A C C T C A A ----- 492
 QY 117 C y s v a l l y s a r g g l y l e u v a l l y s g l u g l y l u a s p p r o g l i n l y e v a l l a l a l a s p g l n 136
 DB 493 ----- G T G A C A C T G A T T A C C A C A A G A T A --- G T T A T A G C 525
 QY 137 l e u v a l l y a s p g l y a l a s p i l e u n l s t h r l l e g l y l y a s p a s p t h r a n t h r a l a 156
 DB 526 A T T C A A G A T C A G A A T C A T C A G G T T A C A T T A T T G A G A G A T G A A C C A G A G A G T 585
 QY 157 A l a l a s p l e u a l a l a p h e u l a l a r g a n a n t y r g l y l e u t h r v a l l i e g l y l e u 176
 DB 586 G C T T A G T A A T A T T T G A G A A T T G A A G A C T C A T A A A G T C T G T G T G G A A T T 645
 QY 177 P r o l y s t h r v a l a s p a s p v a l p h e p r o l l e y s g l n s e r l e u g l y a l a t r p t h r a l a 196

Db	646	CCAAAATACTATGATATGATATATTCGCGGTGATAGTAATATCTTTGGGTTTGACACTGCT	705
Qy	197	AlaGluGlnGlyAlaArgTyrPhe-----MetAsnValValAlaGluAsnAsnAlaAsn	214
Db	706	GTGAGGAAAGCTCAACGAGCGCATTTAAACGACAGACATGTTTGAAGCCGAAAGTAATGACAT	765
Qy	215	ProArgMetLeuIleValHisGluValMetGlyArgAsnGlyGlyTyrPheThrAlaAla	234
Db	766	GGTATCGGTTTGTGTC-----AAGCTTATGGGCTGCTTACAGCGGGTACATACCG-----	813
Qy	235	ThrAlaGlnGluTyrArgIleValLeuLeuAspArgAlaGluTyrPheProGluLeuGlyLeu	254
Db	814	-----ATGATAGCT	822
Qy	255	ThrArgGluSerTyrGluValHisValAlaPheValProGluMetAlaIleLeuPheGlu	274
Db	823	ACATTAGCCACAGACAGATGTTGACTGTGCTGTGATGCCGAGTCCACATTTTACCTCGAA	882
Qy	275	AlaGluAla-----LysArgLeuAspGluValMetAsnLysVal	287
Db	883	GGAGAAGGTGACTCTTTGAGTTCATAGAGAGACGGCTGAAGAGATGTCACATGGTGTG	942
Qy	288	AspCysValAsnIlePheValSerGlnValAlaGlyValAlaIleValAlaGluMet	307
Db	943	-----ATTGCTCTTGCTGTAAGGCTGACAGACAGATTTGATGTGCAAAAGCATG	990
Qy	308	GlnAlaLeuGlyGlnGluValProArgAspAlaPheGlnHisIleLeuLeuAspAlaVal	327
Db	991	GAATCT-----ACTGCTATGGATGGCTTCTGGGAACAAACTTCTTAAGATGTC	1038
Qy	328	AsnProGlyIleTyrPheGlyGlnGlnPheAlaGlnMetIleGlyAlaGluIleThrLeu	347
Db	1039	-----GGCTTGCGCTATCACAAAGCATCAAGATCATTTTAAGAAATTAAGTGTGTG	1092
Qy	348	ValGln-----LysSerGlyTyrPheAlaArgAlaSerAlaSerAsnValAsp	363
Db	1093	ATGAATCTCAAAATACATGATGATCTCTACATACATGATCCGGGCTGTTCCAAAGTAATCATCA	1155
Qy	364	AspMetAspLeuIleLeuSerCysAlaAspLeuAlaValGluCysValAspAspArgGlu	383
Db	1153	GACAAATGTTTACTGTACACTTCTTCTGACAGCGCGGTTCAATGTCGCAATG-----	1203
Qy	384	SerGlyValIleGlyHisAspGluAsp-----AsnGlyAsnValLeuArgAlaIleGlu	401
Db	1204	GCTGGTTACACTGGCTACACACAGTGGTCTTGTCAATGGAGAC--CAAAATACATCCCT	1260
Qy	402	PheProAspIleLeuGlyGlyLeuSerPheAsnIleAspThrAsp-----TyrPheAsn	419
Db	1261	TTCTTACAGAGTAATACAGAGACCCAGAAACATGTAGATTAAGATTAAGATGTGGCGGAG	1320
Qy	420	SerMetLeuSerGluIleGlyGlnPro	428
Db	1321	---CTATTGCTTCTTACGACACGACCA	1344
RESULT 9			
ACA23197			
ID ACA23197 standard; DNA; 1347 BP.			
XX ACA23197;			
XX	19-JUN-2003	(first entry)	
XX	DE	Prokaryotic essential gene #4854.	
XX	DE	Prokaryotic essential gene #4854.	
KM	Antisense; dir. prokaryotic essential gene; cell proliferation;		
KM	drug design; gene.		
XX	Borrelia burgdorferi.		
XX	OS		
XX	W0200277183-A2.		
XX	03-OCT-2002.		
XX			

PF 21-MAR-2002; 2002ZMO-US0091107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XA (EUIT-) EUITRA PHARM INC.
FA
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-023926/02.
DR P-PSDB; ABUI9327.
PT
PT New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 11067; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than S. aureus, S. typhimurium, K. pneumoniae or P. aeruginosa. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1347 BP; 427 A; 169 C; 289 G; 462 T; 0 U; 0 Other;

Alignment Scores:

	Pred. No.:	Score:	Length:
Percent Similarity:	270.00	Matched: 102	1347
Best Local Similarity:	41.71%	Conservative: 69	
Query Match:	24.88%	Mismatches: 177	
	11.82%	Indels: 62	
		Gaps: 11	

US-10-701-200-6 (1-437) x ACA23197 (1-1347)

34 LySValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySer 53
Db 241 AAAGCTGTATTATACACTTGGTGCGCTTTGTCCTGGTTTAAAGATATTATTCCTCT 300
54 LeuilegtuarqYrThGluiLeaspProserllegluileilcCyStyArnglYgly 73
Db 301 ATTGTGCGAATTTATGAAAATATATGAGGGGTCCCAATATTATGAGTAATAATTGGA 360


```

Db      361 TCGGCACTTATCACTGATGGTGTACCAAGTCTTGGAGTTGATGAGAGGTCACAA
Qy      76 G1yleuLeuLeuGlyAspSerTyrrProValThrAlaGluValArgLysValAlaGlyVal
Db      421 GGTTTTACTCAAAAAACACCATCACTTGTGACTCCAAAGTT-----GTAAATGAC
Qy      96 LeuGlnArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThrAsnValLys
Db      472 ATCCATAAAGCGTGTGATCAATTATTTGGACCTTCGT-----510
Qy      116 AspCysValLysArgGlyLeuValLysGlyGlyLysProGlnLysValAlaAsp
Db      511 -----GGGGGCGCATGATTAACCAAGATA--GTTGAC 540
Qy      136 GlnLeuValLysAspGlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThr
Db      541 AGTATTCAGAGATCGTGTGATCAATCAGGTTTATATATGGAGACGCTACTCAAAAG
Qy      156 AlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTyrrGlyLeuThrValIleGly
Db      601 GGAGCAGCTGTATTATTCAGAGAGTGAAGCGGTGAACCTTAACCTGATGCTGGG 660
Qy      176 LeuProLysThrValAspAsnAspValPheProIleLysGlnSerLeuGlyAlaTrpThr
Db      661 ATTCCTAAGACATGATGATATGACATTCGCGTCATGATGAAGCTTTGGTTTGACACG
Qy      196 AlaAlaGlnGlnGlyAlaArgTyrrPheMetAsnValValAlaGlnLysAsnAlaAsnPro
Db      721 GCTGTGGAAAGGCTCAAGCTCCATTAAATGCTGCACATGTGGAGGCTGAAGTGTGAG
Qy      216 ArgMetLeuIleValHisGluValMetGlyArgAsnCysGlyTrpLeuThrAlaAlaThr
Db      781 AATGGCATAGGGGTGTCAACTATGAGGACGCTAATAGGANTCATCCGA-----831
Qy      236 AlaGlnGlyTyrrArgLysLeuLeuAspArgAlaGluTrpLeuProGlnLeuGlyLeuThr
Db      832 -----ATGATGCACT 843
Qy      256 ArgGlnSerTyrrGluValHisAlaValPheValProGlnMetAlaIleLeuLeuGlyAla
Db      844 TTGGCTAGTCAGATGTGATTTATGTTATTCCTGAATACCTTTTATCTTGGAGGA 903
Qy      276 GluAla-----LysArgLeuArgGluValLysMetAspLysValAsp 288
Db      904 GAAAGCGCATTTTAAATATATGAGAAAAAGCTCTCAAGACATGAGACATAGCTC---960
Qy      289 CysValAsnIlePheValSerGlyValAlaGlyValAlaIleValAlaGlnMetGln
Db      961 -----ATCGTGTACAGAAAGCTGTGTGACGAGCTGCTCAGAGAAACTTG 1011
Qy      309 AlaLysGlyGlnGluValProArgAspAlaPheGlyHisIleLysLeuAspAlaValAsn
Db      1012 AAA-----ACTTCAACCGCAAAAGATGCTTCTGAAATTAACCTTTCACAGATGC---1062
Qy      329 ProGlyLysTrpPheGlyGlnGln-----PheAlaGlnMetIleGlyAlaGln 344
Db      1063 ---GGATTGTGATTTCTGATAGATTAAAGCTCATCTTCTTAAATTCCTCCATGCC 1119
Qy      345 LysThrLeuValGln---LysSerGlyTyrrPheAlaArgAlaSerAlaSerAsnValAsp 363
Db      1120 ATTACTCTCAAAATACATAGATCCAACTTACATGATCCGTGCGGTCCAAATATGATGCT 1179
Qy      364 AspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArgAlaGln 383
Db      1180 GATTAAT-----GTATAGTGCACCTCTCTTGTCTCAAGTGTCTTACAGAGTATG 1230
Qy      384 SerGlyValIleGlyHisAspGluAsp-----AsnGlyAsnValLeuArgAlaIleGln 401
Db      1231 GCGGCGTACACCGGCTTTCACAACTGGGCTTGTCAATGCTGA---CAGACTTATATTCCA 1287
Qy      402 PheProArgIleLysGlyGlyLysProPheAsnIleAspThrAsp-----TrpPheAsn 419

```

```

Db      1288 TTTAAATGATATCACTGAGACAGCAATTAACGTTGTGATTAACGATGCTGGGACAG 1347
Qy      420 SerMetLeuSerGluIleGlyGlnPro 428
Db      1348 ---CTTCTGTATCCACCAACCAACCA 1371

RESULT 11
AXX20248_01/c
Continuation (2 of 10) of AXX20248 from base 100001 (Borrelia burgdorferi polynucleotide
WP Sequence split into 10 fragments LOCUS AXX20248 Accession Axx20248
WP Fragment Name Begin End
WP AXX20248_00 1 110000
WP AXX20248_01 100001 210000
WP AXX20248_02 200001 310000
WP AXX20248_03 300001 410000
WP AXX20248_04 400001 510000
WP AXX20248_05 500001 610000
WP AXX20248_06 600001 710000
WP AXX20248_07 700001 810000
WP AXX20248_08 800001 910000
WP AXX20248_09 900001 910715

Alignment Scores:
Pred. No.: 1,48e-15 Length: 110000
Score: 270.00 Matches: 102
Percent Similarity: 41.71% Conservative: 69
Best Local Similarity: 24.88% Mismatches: 177
Query Match: 11.82% Indels: 62
DB: 2 Gaps: 11

US-10-701-200-6 (1-437) x AXX20248_01 (1-110000)
Qy      34 LysValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySer 53
Db      45111 AAGCGCTATTACATCTGTGTGGCTTTGCTTCGTTTAAACATGTTATTCCTCT 45052
Qy      54 LeuIleGluArgTyrrThrGluIleAspProSerIleGluIleLysCysTyrrArgGlyGly 73
Db      45051 ATTGTGCAAACTTTATGAAATTAATATGAGGTTCGCAATATTATTAAGCTAAATTTGGA 44992
Qy      74 TyrLysGlyLeuLeuLeuGlyAspSerTyrrProValThrAlaGluValArgLysValAla 93
Db      44991 TATCAAGGCTTCTTCTGAGTCAAAATTCACCTTTTATTAATCTTAATCCAGATGTTGT 44932
Qy      94 GluValLeuGlnArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThrAsn 113
Db      44931 GATGATATTAAATTAATTTGAGAGTACTATTCTTGTTCTCA-----44890
Qy      114 ValLysAspCysValLysArgGlyLeuValLysGlyGlyLysAspProGlnLysValAla 133
Db      44889 -----AGGGGTGTGATTAAGCCGTGGAA-----ATA 44863
Qy      134 AlaAspGlyLeuValLysAspGlyValAspIleLeuHisThrIleGlyGlyAspAspThr 153
Db      44862 GTTGATACCTTTAGAGAAATGAATGAACTTAATGATTTTAAACATTTGCGGGAGTGTACT 44803
Qy      154 AsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTyrrGlyLeuThrVal 173
Db      44802 CAAAAGGCTCTCTTTATTTGCTGAGACATGAAGAAAAAGAAATTTAAATAATGACAGTT 44743
Qy      174 IleGlyLeuProLysThrValAspAspValPheProIleLysGlnSerLeuGlyAla 193
Db      44742 GTGGGCACTTCCTTAACCGTAAAGCATATGATTTTATGTTTTCACAAAATCTTTGGATTT 44683
Qy      194 TrpThrAlaAlaGlnGlnGlyAlaArgTyrrPheMetAsnValValAlaGlnLysAsnAla 213
Db      44682 GAGACTGCTGTGAACACAGCTGTGCACTGTGCTGCTCATTTTGAAGCTAATAGT 44623
Qy      214 AsnProArgMetLeuIleValHisGluValMetGlyArgAsnCysGlyTrpLeuThrAla 233
Db      44622 GCTTATTAATGCACTTGTGATTTAAAGTTATGGAGCGGGAATTCGCGCTTTATGCTGCT 44563
Qy      234 AlaThrAlaGlnGlyTyrrArgLysLeuLeuAspArgAlaGluTrpLeuProGlnLeuGly 253

```

```

Db      44562 CTTACGCA----- 44554
Qy      254 LeuThrArgLysSerGluValHisAlaValPheValProGluMetAlaIleAspLeu 273
Db      44553 -----CTTCTCTTAATGATGTTAATTTTGTATATCCCAAGAACTGATTTGACATA 44500
Qy      274 GluAlaGlu-----AlaLysArgLeu-----ArgGluValMet 284
Db      44499 GAAGGCTCTAATGATTTCTTGTTCATCTTGAAGACGCTTTAGAGAAAGAAAGCTTAA 44440
Qy      285 AspLysValAspCysValAsnIlePheValSerGluGluAlaGluValGluAlaIleVal 304
Db      44439 GAAGGATTCCTCCATGCACTAATTTGATAGCAAGAGGCGAGC----- 44395
Qy      305 AlaGluMetGluAlaLysGluGluGluValPro-----ArgAspAlaPheGluHisIle 322
Db      44394 -----CAAAATATTTTGTATCATTTTCCTAAGAAAGAAAGACGATTCGGGAATTTG 44344
Qy      323 LysLeuAspAlaValAsnProGluLysTrpPheGluGluGlnPheAlaGlnMetIleGly 342
Db      44343 CTTTATGAGATAT-----GGGCTTTATATTAAGATAAATTAACGAAATATTTTAA 44290
Qy      343 AlaGluLys-----ThrLeuValGln-----LysSerGlyTrpPheAlaArgAlaSer 358
Db      44288 GCAAAATAATATCAATTTACTTTAAATACATTTGATCCAGCTAATATTTAGAAAGTTCA 44230
Qy      359 AlaSerAsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCys 378
Db      44229 CCTGCTAATGCCAGATTCGCTTATTTGTGCTAGCGCTTGGGTCGAAATCTGTGCTGT 44170
Qy      379 AlaPheArgArgLysSerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArg 398
Db      44169 GCAATGCGCTGTAGACAAATAATGTTAGTTAGTTATGAGTACAAATAATTTGTCATATA 44110
Qy      399 AlaIleGluPheProArgIleLysGlyLysProPheAsnIleAspThrAspTrpPhe 418
Db      44109 CCTATTATAGATGCGATTAATTTGACAGAAATTAAG--GTTAATCCAAATGCTCTTTTGG 44053
Qy      419 AsnSerMetLeuSerGluIleGlyGlnPro 428
Db      44052 AGAGATGTTCTTTCAAGTACGGGGGACCA 44023
RESULT 12
AA085984
ID      AA085984 standard; cDNA to mRNA; 1624 BP.
AC      AA085984;
XX
DT      25-MAR-2003 (revised)
DT      12-OCT-1995 (first entry)
XX
DE      Oryza sativa pPFK-OS1 gene.
XX
KM      ATP-dependent fructose-6-phosphate 1-phosphotransferase enzyme; plant;
KM      potato; Solanum tuberosum; rice; Oryza sativa; maize; Zea mays; radish;
KM      Rapihanus sativus; Flaveria brownii; primer; expression vector;
KM      Agrobacterium tumefaciens; sugar; storage; temperature; ds.
XX
OS      Oryza sativa.
XX
FH      Key Location/Qualifiers
FT      CDS 3..1412
FT      /tag= a
FT      /product= "fructose-6-phosphate 1-phosphotransferase"
FT      /tag= b
FT      /transl_except= seq: AAT, a.a.:11e
XX
PN      W09505457-A1.
XX
PD      23-FEB-1995.
XX

```

```

PF      16-AUG-1994; 94MO-JP001352.
XX
PR      19-AUG-1993; 93JP-00226454.
XX
PA      (NIBS) JAPAN TOBACCO INC.
XX
PI      Hiroyoshi T, Mine T, Kasaoka K, Tyson HR, Page MJ;
XX      WPI, 1995-098757/13.
XX      P-PEDB; AAR71581.
XX
PT      DNA coding for fructose-6-phosphate 1-phosphotransferase - of plant
PT      origin, for prodn. of transformant plant cells with altered sugar
PT      content.
XX
PS      Claim 8, page 46-49; 79pp; Japanese.
XX
CC      The sequences (AA085982-6) represent the genes encoding a novel ATP-
CC      dependent fructose-6-phosphate 1-phosphotransferase enzyme (EC 2.7.1.11;
CC      PFK) from a range of plants. This is the sequence of the Oryza sativa
CC      (rice) gene, pPFK-OS1, as given in the specification. Plants transformed
CC      with these genes can express the enzyme. The transformed plants can
CC      produce varieties that have altered sugar content on storage at low
CC      temperatures. (Updated on 25-MAR-2003 to correct EN field.)
XX
SQ      Sequence 1624 BP; 480 A; 302 C; 406 G; 436 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,74e-17 Length: 1624
Score: 264.50 Matches: 109
Percent Similarity: 40.54% Conservative: 56
Best Local Similarity: 26.78% Mismatches: 173
Query Match: 11.58% Indels: 69
DB: 2 Gaps: 16

US-10-701-200-6 (1-437) x AA085984 (1-1624)
Qy      37 ILeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGlu 56
Db      222 ATTGCACATGTGGAGGACTGTGCTGCTGACTGAACACTGCATATTAAGCAATTTGTTCT 281
Qy      57 ArgTrpThrGluIleAspProSerIleGluIleIleCysTrpArgGlyGlyTrpLysGly 76
Db      282 GGCTTAATAGCATATATGATGTGTCAGTAGGACTTGAATTCAGGAGGTGATAGAGT 341
Qy      77 LeuLeuLeuGlyAspSerTrpProValThrAlaGluValArgLysLeuAlaValLeu 96
Db      342 TTCTATGCTTGAACACATGACTGACTGATCA-----AGAGGTAAACAGCAAT 392
Qy      97 GlnArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAsp 116
Db      393 CACAAAGGGGTGGAACGTTCTTGGGACATCAGT----- 428
Qy      117 CysValLysArgGlyLeuValLysGluGluAspProGlnLysValAlaAlaAspGln 136
Db      429 -----GGAGGCCATGACACCATGAAGATT--GTTGACAGC 461
Qy      137 LeuValLysAspGlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAla 156
Db      462 ATCCAGATCGGTGATTAATACAGTTTATGTAATTTGTTGATGATGATCAAGAGGCT 521
Qy      157 AlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTrpGlyLeuThrValIleGlyLeu 176
Db      522 GCAGAGGATTTTGAAGATTAAGAAGACGTGCTCAAGGTTGCTGCTGCTGCTGCAAT 581
Qy      177 ProLysThrValAspAspValPheProIleLysGlnSerLeuGlyAlaIleTrpThrAla 196
Db      582 CCAAGACGATTAATGATATATACAGTAATTCACATATTTGTTGTTGCACTGCA 641
Qy      197 AlaGluGlnGlyAlaArgTrpPheMetAsnValValAlaGluAsnAlaAsnProArg 216
Db      642 GTTAGAGGCGCCAAAGTCATTAATAGTCTCATGTAGAAAGCTGGAAGCCCGAGAAAT 701

```

```

QY 217 MetLeuIleValHsegluValMetGlyAAsnCysGlyTyrLeuThrAlaIleThrAla 236
DB 702 GGTATAGGCGCTCTTAAGTAATGGGTGCGACAGTGGTTTATGTCA----- 749
QY 237 GlnGlyTyrArglyLeuLeuAspArgAlaGlyTyrLeuProGluLeuGlyLeuThrArg 256
DB 750 ---CACTAAGCTACTCTAGCCACAG----- 773
QY 257 GluSerTyrGluValHsegluValPheValProGluMetAlaIleAspLeuGluAlaGly 276
DB 774 -----GACGTGATGTTGTTGATTCCAGAGTCAAGCTTTTCAATCTCGAAGGTGA 824
QY 277 Ala----- 878
DB 825 GGTGGCTTTTATGATATTTTGAAGGCGTCTAGAGAGATGTCATATAGTT----- 878
QY 290 ValAsnIlePheValSerGlyGlyAlaGlyValGluAlaIleValAlaGluMetGlnAla 309
DB 879 -----ATGCTTTTCGGAGGTGCGGCGAGAACTTTATTAAGAAACAAAGATCA 932
QY 310 LyseGlyGlnGluValProArgAspAlaPheGlyHis---IleLyseLeuAspAlaValAsn 328
DB 933 ATGGGG-----AAAGATGCTTCAGGCAATTCATCTTGTATGTT----- 974
QY 329 ProGlyLyseTyrPheGlyGluGlnPheAlaGluMetIleGlyValaGlyLeuThrLeuVal 348
DB 975 ---GGTCTTGTGTTATCTCAAAAGATAAAGACATTTCAAGAAATCAAGACTACTATA 1031
QY 349 Gln----- 1091
DB 1032 AATCTCAAGTATATTAATCTTACATACATGATGCGCCATTCCTGCAATGCAATCTGAC 1091
QY 365 MetArgLeuIleLyseSerCysAlaAspLeuAlaValGluCysAlaPheArgCysGlySer 384
DB 1092 AATGTATATTCACACTGTTGGCACACAGGCTGTTTCATGAGACCCTAGGTGATACACT 1151
QY 385 Gly---ValIleGlyHisAspGlyAspAsnGlyAsnValLeuArgAlaIleGluPhePro 403
DB 1152 GGTTCACACTGTTGGC-----CAAGTAAATGTCGCGCATTCG---TATATCCCGTTTAC 1202
QY 404 ArgIle-----LyseGlyLyseProPheAsnIleAspThrAspTyrPheAsnSerMet 421
DB 1203 AGGATACAGAGAGAGAAAGAAAGTTTCAATTACTGATGATGTCGCGCAGA---CTT 1259
QY 422 LeuSerGluIleGlyGlnPro 428
DB 1260 CTCTCTCAACCAACGACCA 1280

RESULT 13
ADA70655
ID ADA70655 standard; DNA; 2256 BP.
XX
XX ADA70655;
XX
XX 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 3978.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; db.
XX
XX Oryza sativa.
XX
XX MO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001MO-IB001105.
XX
XX 22-JUN-2001; 2001MO-IB001105.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX

```

```

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
PI Katagiri F, Qian S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
DR WPI; 2003-175290/17.
XX
XX
XX PT Identifying at least one gene involved in plant resistance or response to
XX pathogenic infection for conferring resistance or tolerance to a plant to
XX PT bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
PS Claim 6; SEQ ID NO 3978; 899bp; English.
XX
XX
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
SQ Sequence 2256 BP; 568 A; 574 C; 665 G; 449 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3, 47e-17 Length: 2256
Score: 263.50 Matches: 96
Percent Similarity: 40.71% Conservative: 53
Best Local Similarity: 26.23% Mismatches: 158
Query Match: 11.54% Indels: 59
DB: 8 Gaps: 9

US-10-701-200-6 (1-437) x ADA70655 (1-2256)
QY 34 LyseValAlaIleLeuThrAlaGlyIleValaProCysLeuAsnSerAlaIleGlySer 53
DB 487 AAAGCATGATGTTTAACTCGCGAGGCGCTTGGCTTAAATACGGTCATTAAGAG 546
QY 54 LeuIleGluArgTyrThrGluIleAspProSerIleGluIleIleCysTyrArgGlyGly 73
DB 547 TTGGTGTGGCGCTTGGCCCACTGACATGACATGATGATCTTCCATGACACCT----- 606
QY 74 TyrLyseGlyLeuLeuLeuGlyAspSerTyrProValThrAlaGluValArgLyseVala 93
DB 607 TACAGCGATTTCTATTCAGATATTAATCTTCCATGACACCT-----AAAAGTGC 657
QY 94 GlyValLeuGlnArgPheGlyGlySerValIleGlyAsnSerArgValLyseLeuThrAsn 113
DB 658 AATGATATCCACAAAGAGGTGCGACAGTTCGGAACATCACGG----- 702
QY 114 ValLyseAspCysValLyseArgGlyLeuValLyseGluGlyAspProGluIlyseVala 133
DB 703 -----GGTGTCTATGACACCAAAAATTT--- 726
QY 134 AlaAspGlnLeuValLyseAspGlyValaAspIleLeuHisThrIleGlyIlyAspAspThr 153
DB 727 GTTGACACATTCACAAATGCTGTTATTAATGAGTATGATGAGCAAGGCGATGGAAC 786
QY 154 AsnThrAlaAlaIleAspLeuAlaIlePheLeuAlaArgAsnAsnTyrGlyLeuThrVal 173
DB 787 CAGAGCGAGCATATGAAATATATAGAAATTCGTAGACGTGCTAAAGTTCGTGTT 846
QY 174 IleGlyLeuProLyseThrValaAspAsnAspValaPheProIleLyseIleSerLeuGlyAla 193
DB 847 GCTGTGTTCCCAAGCAATGATATGATATGATAGCGGTTATGACAAAGTCTTTGGTTT 906
QY 194 TrpThrAlaAlaGluGlnGlyAlaArgTyrPheMetAsnValaIleGluLeuAsnAla 213
DB 907 GATAGTGTCTAGAAAGAGCCAGCTGCATTTGATGACAGCTCATCTTGAAGCTTCAAG 966
QY 214 AsnProArgMetLeuIleValHsegluValMetGlyAAsnCysGlyTyrLeuThrAla 233

```

```
Db 967 GGTGAATGGAATAGCTTACTTAACATGATGGGCTATAGTCGGTTATTGCA--- 1023
Oy 234 AATHTTAAAGLGLuTyrArgLysLeuLeuAspArgAlaGLuTrpLeuProGluLeuGly 253
Db 1024 -----ATGTAAT 1029
Oy 254 LeuThrArgGluSerTyrGluValAlaAlaValAlaPheValProGluMetAlaIleAspLeu 273
Db 1030 GCTACTCTTGAAGACAGACAGCTGACTGCTTAATTCCTGAAGTACACCATTTATTTGG 1089
Oy 274 GUAAGLUALA-----LysArgLeuArgLysValMetAspLys 286
Db 1090 GAAGGGAGGGGTGAGACTGTTTGAATACATAGAGAAAGGTTGAAGACAAACACACATG 1149
Oy 287 VALAspCysValAsnIlePheValSerGluGlyValGlyValGluAlaIleValAlaGlu 306
Db 1150 GTT-----ATTGTTGAGCTGAGGAGACAGACAGATCTTATTTGCCAAAGT 1197
Oy 307 MetGlnAlaLysGlyGlnGluValProArgAspAlaPheGlyValIleLysLeuAspAla 326
Db 1198 ATAGCTGCAGCATCA-----ATAGATGCATCTGAAATTAAGCTACTT----- 1242
Oy 327 VALAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMetIleGlyValGluLysThr 346
Db 1243 CTTGATCTTGGTCTTGGCTGACTCAACAAGATTAAAGATTACTGCACAAACAAAAGATG 1302
Oy 347 LeuValGln-----LysSerGlyTyrPheAlaArgAlaSerAlaSerAsnVal 362
Db 1303 GAGATGACTATTAAATATATAGTATAGTCCAACTCAATGATCCGTCATTCACAGCATGCA 1362
Oy 363 AspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArgArg 382
Db 1363 TCAGCAATGTCTACTGCACTGCTGCGCACAGTGCATTCATGAGGCAATGCGCTGGA 1422
Oy 383 GluSerGlyValIleGly 388
Db 1423 TATAGCTTCACAGTTGCA 1440

RESULT 14
AA085986
ID AA085986 standard; cDNA to mRNA; 1558 BP.
XX
AC AA085986;
XX
DT 25-MAR-2003 (revised)
DT 12-OCT-1995 (first entry)
XX
DE Raphanus sativus pPFK-ZM1 gene.
XX
KW ATP-dependent fructose-6-phosphate 1-phosphotransferase enzyme; plant;
KW potato; Solanum tuberosum; rice; Oryza sativa; maize; Zea mays; radish;
KW Raphanus sativus; Flaveria brownii; primer; expression vector;
KW Agrobacterium tumefaciens; sugar; storage; temperature; de.
XX
OS Raphanus sativus.
XX
FH Key Location/Qualifiers
FT CDS 1..1281
FT /tag= a
FT /product= "fructose-6-phosphate 1-phosphotransferase"
XX
FN W09505457-A1.
XX
PD 23-FEB-1995.
XX
PF 16-AUG-1994; 94WO-JP001352.
XX
PR 19-AUG-1993; 93JP-00226454.
XX
PA (NISB ) JAPAN TOBACCO INC.
XX
PI Hiroyoshi T, Mine T, Kaseoka K, Tyson HR, Page MJA;
```

```
DR MPI.1995-098757/13.
XX P-PSDB; AAR71583.
XX
PT DNA coding for fructose-6-phosphate 1-phosphotransferase - of plant
PT origin, for prodn. of transformant plant cells with altered sugar
PT content.
XX
PS Claim 12, Page 58-61; 79pp; Japanese.
XX
CC The sequences (AA085982-6) represent the genes encoding a novel ATP-
CC dependent fructose-6-phosphate 1-phosphotransferase enzyme (EC 2.7.1.11;
CC PFK) from a range of plants. This is the sequence of the Raphanus sativus
CC (radish) gene, pPFK-RS1. Plants transformed with these genes can express
CC the enzyme. The transformed plants can produce varieties that have
CC altered sugar content or storage at low temperatures. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
SQ Sequence 1558 BP; 470 A; 280 C; 373 G; 435 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,36e-17 Length: 1558
Score: 263.00 Matches: 108
Percent Similarity: 41.39% Conservative: 65
Best Local Similarity: 25.84% Mismatches: 173
Query Match: 11.51% Indels: 72
DB: Gaps: 17

US-10-701-200-6 (1-437) x AA085986 (1-1558)
Oy 37 ILeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGlu 56
Db 148 ATGTTTCAATGTTGGGCTTGTGTCAGGCTTAATCTGATCTGATAGAAATCTTTGT 207
Oy 57 ArgTyrThrGluIleAspProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGly 76
Db 208 GGAATGTCTTAATCATGTATGATGTCAGAAATCTTGCCATTCAGAGAGGATTACAGAGGC 267
Oy 77 LeuLeuLeuGlyAspSerTyrProValThrAlaGluValArgLysValAlaGluLeu 96
Db 268 TTCTAGCTTAGAACAACATGATTCGATTGATTG-----AAAACAGTAATGATATT 318
Oy 97 GlnArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAsp 116
Db 319 CATAAACCTGAGGAACCATCTCGGACTTCAGCA----- 354
Oy 117 CysValLysArgGlyLeuValLysGluGlyLysAspProGlnLysValAlaAlaAspGln 136
Db 355 -----GTCGTCACGACACTACTAAGATA---GTTGATAGT 387
Oy 137 LeuValLysAspGlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAla 156
Db 388 ATTCAAAGATCGTGGGATTAAACAGGTTTATTAATTCGGTGGAGATGCACAGAAAGGA 447
Oy 157 AlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeu 176
Db 448 GCAGCTGTTATATTCGAGAGATTAGAGACGCTGATCAACAGCTCATGTGGAAGCTTATGAGAT 507
Oy 177 ProLysThrValAspAsnAspValPheProIleLysGlnSerLeuGlyAlaIleTrpThrAla 196
Db 508 CCCAAAACATCGACATGACATTCCTTATTCATATACATGCTTCGGCTTTGACACAGCT 567
Oy 197 AlaGluGlnGlyAlaArgTyrPheMetAsnValValAlaGluAsnAlaAsnProArg 216
Db 568 GTAGAAAGAGCTCAACGCTATCAACACAGCTCATGTGGAAGCTTATGAGAT 627
Oy 217 MetLeuIleValAlaGluValMetClyArgAsnGlyLysTrpLeuThrAlaAlaThrAla 236
Db 628 GGTATGTGCTTGTCAAGTTAATGACGTTATAGTGAATTCATTCGCGATG----- 678
Oy 237 GlnGluTyrArgLysLeuLeuAspArgAlaGluTrpLeuProGluLeuGlyLeuThrArg 256
Db 679 -----TATGCAAGCTAGCCACGACGA----- 699
```


Qy	257	GluseYrYgluValHhlaValPheValProglumetalalaePhegluValGlu	276
Db	700	-----GACGTGAGACTGCTGCTGATCCCGGAAATCCATTTTCTTGAAGGCAA	750
Qy	277	-----AlaValargLeuArgGluValMetAspLysValaspCys	289
Db	751	GGCGGTCTTTTCGAGTTTATCGGTAAACGGCTTAAAGAGATTGGTCATGTG-----	804
Qy	290	ValenIIlePheValSerGluGlyAlaGlyValGluAlaIleValAlaGluMetGlnAla	309
Db	805	-----ATTGATGATAGCAGAAAGTGCTGGA---CAAGATCTGTTGGCTGAAAGCAATGAA	855
Qy	310	LysGlyGlnGluValProArgAspAlaPheGlyHisIleLysIleLysLeuAspAlaValAsnPro	329
Db	856	CAGTCCACCAACCTTC---AAAGATGATCTGGGAAACAACTTTACAGACGTT-----	906
Qy	330	GlyIysTrpPheGlyGluGlnPheAlaGlnMetIleGlyAlaGluLysThrLeuValGln	349
Db	907	GGCCTATGATCTCCCAACGGATCAAGATCATTTTGGCAAGAAATACCTTAACTG	966
Qy	350	Lys-----SerGlyTrpPheAlaArgAlaSerAlaSerAsnValaspPheArg	366
Db	967	AAATATCATATAGATCCAACTCATATGATTAAGGCTGTTCGACCAATGATCATGAC	1020
Qy	367	LeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArgArgGluSerGlyAla	386
Db	1021	-----AATGTATGCTGCACCGCTGTAGCTCAACACGGGGTTCAT-----GGAGTGG	1065
Qy	387	Ile---GlyHleAspGluAspAsnLysAsnValLeuArgAla-----IleGlu	401
Db	1066	ATGCTGTATACAAATGAGCTTCAACGCTTGCTCTTGTAAATGGCAGACATATTACATTCC	1125
Qy	402	PheProArgIle-----LysGlyGlyLysProPheAsnIleAspThrAspTrpPheAsn	419
Db	1126	TTCTATAGATCATCTGAGAAACAGAAACAAGGTGGTATCAGTACAGATGGGCAAGG	1185
Qy	420	SerMetLeuSerGluIleGlyGlnProLysGlyLysValGluValSerHis	437
Db	1186	---CTTTTGTCTTCGACAAACGAGCTGATTTCATGAAGCAGATGATCACACC	1236
RESULT 15			
ID	ADS48162	standard; cDNA, 960 BP.	
XX	ADS48162;		
AC			
XX			
DT	02-DEC-2004	(first entry)	
XX			
DE	Bacterial polynucleotide #2905.		
XX			
KW	Recombinant DNA construct; transformed plant; improved plant property;		
KW	cold tolerance; heat tolerance; drought tolerance; herbicide; omeosis;		
KW	pathogen tolerance; pest tolerance; plant disease resistance;		
KW	cell cycle pathway modification; plant growth regulator;		
KW	homologous recombination; seed oil yield; protein yield; carbohydrate;		
KW	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;		
KW	bacterial polynucleotide; gene; ss.		
XX			
OS	Bacteria.		
XX			
PN	US2003233675-A1.		
XX			
PD	18-DEC-2003.		
XX			
PE	20-FEB-2003; 2003US-00369493.		
XX			
PR	21-FEB-2002; 2002US-0360039P.		
XX			
PA	(CAOY/) CAO Y.		
PA	(HINK/) HINKLE G J.		
PA	(SLAT/) SLATER S C.		
PA	(CHEN/) CHEN X.		
PA	(GOLD/) GOLDMAN B S.		

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 DR WPI, 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PT
 XX
 PS Claim 1, SEQ ID NO 26592, 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 CC
 XX
 SEQ Sequence 960 BP; 295 A; 196 C; 278 G; 191 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,55e-17	Length:	960
Score:	262.00	Matches:	90
Percent Similarity:	47.15%	Conservative:	59
Best Local Similarity:	28.48%	Mismatches:	113
Query Match:	11.47%	Indels:	55
DB:	13	Gaps:	14

US-10-701-200-6 (1-437) x ADS48162 (1-960)

QY	33	LyblybValAlAlleleuthrAlaGlYglYleuAlaProCyLeuAmsenrAlaIleGly	52
DB	4	AAAGAAATAGGACGTACTTACAAACGGCGGAGACGGCACTCGAATGAACCAACGACTGGAGA	63
QY	53	SeerleuIleGluArGTyrThrGluIleAspProserIleGluIleIleCysTyrArGly	72
DB	64	GCGCGGTC--AGGTAC--GCTGCAGGACGAGCACTGAAAGTATTCGAGTGAAGA	117
QY	73	GlYTrYbGlYleuLeuGluYAspSerTyrProValThrAlaGluValAspLys	92
DB	118	GGTTCCTCAGCGCTCATGACGGCATT--GTAACCTCGATACAAAGATGTG	171
QY	93	AlaGluValLeuGluArGpHeGlYglYserValIleGlyAmsenrArGValYbLeuThr	112
DB	172	GCAGAAATCACAAGAAAG--GAGAGAACAAATTCGAGAACTTCCAGATGTGGAGGTTTC	238
QY	113	AenVallybAspCyVallybArGlyLeuVallybGluGlybAspProGlnlybVal	132
DB	229	AAAGAC-----GACAGGGCAGGAACTC	252
QY	133	AlaAlaAspGlnLeuVallybAspGlyValaAspIleLeuHisThrIleGlyglYAspAsp	152
DB	253	GCTGGAAACAGATATAAAAACATGCTATTAAGAAAGCTCGTCTCATAGGTGTGAAGG	312
QY	153	ThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArGAsnAsnTyrGlYleuThr	172

Db 313 AGCTCAAC-----GGCGCTCATCTTCTTTACGAAGAACACAAATGCC 357

Qy 173 ValIleGlyLeuProLysThrValAspAsnAspValPheProIleLysGlnSerLeuGly 1922

Db 358 GTTGCGGTATACCAACAACCATAGACACAGCATTTGGGTTGACTGACACATGTGATGGG 417

Qy 193 AlaTrpThrAlaIleGlnGlnGlyAlaArgLysPheMetAsnValAlaGluAsnAsn 2122

Db 418 GTGAGACACGTTTGAACAACGGGTGATGATGCTGTTCAAAAGCTCAAAAGACACCGTTAGC 4777

Qy 213 AlaAsnProArgMetLeuIleValHisGluValMetGlyValGlyAsnGlyTrpLeuThr 2222

Db 478 TCCGATGAGAGACTTTCATTGTGG--GAAGTCATGGGAGAGACATTCGGGTATACGTCT 5344

Qy 233 AlaAlaThrAlaGlnGluLysArgLysLeuAspArgAlaGluTrpLeuProGluLeu 2522

Db 535 CTCATGGCGGGA-----CTGGGAGCTGTTGCAGAA-----5644

Qy 253 GlyLeuThrArgGlnSerLysTrpGluValHisAlaValPheValProGluMetAlaIleAsp 2722

Db 565 -----GCCATCATCTGATACAGAGATTCGCGTGGAT 5944

Qy 273 LeuGluAlaGluAlaLysArgLeuArgGluVal-----MetAspLysValAspCysVal 2900

Db 595 TATTCCACAGCTGCGCATGAAATTCCTCGAAGAAAGAAAGAAAGAAAGATTCACACGATA 6544

Qy 291 AsnIlePheValSerGlnGlyAlaGlyValGluAlaIleValAla--GluMetGlnAla 3090

Db 655 ---ATCATAGTCCGTGAAGGGCAGCGAGTGCCTTACCGTCGCAAGACCTCGAATAC 7111

Qy 310 Lys---GlyGlnGluValProArgAspAlaPheGlyHisIleLysLeuAspAlaValAsn 3288

Db 712 AGGATAGGCTACCAAAACGAGATCAACATCTCGACACGTCACAG-----7588

Qy 329 ProGlyLysTrpPheGlyGlnGlnPheAlaGlnMetIleGlyAlaGlu 3444

Db 759 -----AGGTGGTTCTCCACAGCGCTTTCGACAGAAAGACTGCGACTGTAG 8000

Search completed: March 1, 2005, 15:19:45
Job time : 776 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 1, 2005, 13:55:46 ; Search time 235 Seconds
(without alignments)
3042.778 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284

Sequence: 1 DVTWPHYLTADIRFCMWF.....FNSMLSEIQPKGKVENSH 437

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 segs, 818139359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n model -DB=var1h
-Q/cgn2_1/USFTO_spool/US10701200/runat_25022005_102833_10340/app_query.fasta_1.583
-DB=Issued_Patente NA -OFMT=fasta -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human4.0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFM=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10701200 @CNC 1.1 69 @runat_25022005_102833_10340 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEDUTERY -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued Patente NA:*
- 1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
 - 2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
 - 3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
 - 4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
 - 5: /cgn2_6/prodata/1/ina/PCUTS COMB.seq:*
 - 6: /cgn2_6/prodata/1/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2284	100.0	1311	4	US-09-934-901-15
2	2284	100.0	1311	4	US-09-934-868-5
3	2284	100.0	1311	4	US-10-321-210-15
4	2284	100.0	1311	4	US-10-320-874-15
5	270	11.8	1778	1	US-08-416-870C-3
6	264.5	11.6	1624	1	US-08-416-870C-5
7	262	11.5	1978	1	US-08-416-870C-1
8	261	11.4	1038	4	US-09-902-540-8974
9	261	11.4	10391	4	US-09-902-540-958
10	260	11.4	1558	1	US-08-416-870C-9
11	246.5	10.8	2964	4	US-09-614-221A-404
12	239.5	10.5	6196	4	US-09-902-540-788

13	228	10.0	7972	4	US-08-956-171E-312	Sequence 312, App
14	228	10.0	7972	4	US-08-781-986A-312	Sequence 312, App
15	227	9.9	1017	4	US-09-489-039A-5087	Sequence 5087, App
16	226.5	9.9	1038	3	US-09-134-001C-937	Sequence 937, App
17	226.5	9.9	4244	4	US-09-710-279-4256	Sequence 4256, App
18	223	9.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
19	223	9.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
20	222.5	9.7	1008	4	US-09-583-110-2438	Sequence 2438, Ap
21	222.5	9.7	1086	4	US-09-902-540-7871	Sequence 7871, Ap
22	220	9.6	993	4	US-09-436-681A-4165	Sequence 4165, Ap
23	218.5	9.6	966	4	US-09-134-000C-1927	Sequence 1927, Ap
24	217.5	9.5	19446	3	US-08-961-527-51	Sequence 51, Appl
25	215.5	9.4	972	4	US-09-107-532A-2816	Sequence 2816, App
26	212.5	9.3	4605	3	US-09-221-017B-128	Sequence 128, App
27	211.5	9.3	2048	1	US-08-416-870C-7	Sequence 7, Appli
28	210.5	9.2	640681	4	US-09-790-988-1	Sequence 1, Appli
29	210	9.2	2510	4	US-09-949-016-3617	Sequence 3617, Ap
30	208.5	9.1	2880	4	US-09-614-221A-354	Sequence 354, App
31	206.5	9.0	3047	1	US-08-280-690-1	Sequence 1, Appli
32	204.5	9.0	2591	4	US-09-023-655-1414	Sequence 1414, Ap
33	204.5	9.0	2591	4	US-09-814-915A-102	Sequence 102, App
34	204.5	9.0	2591	4	US-09-949-016-380	Sequence 380, App
35	203.5	8.9	1274	4	US-09-887-054-1	Sequence 1, Appli
36	202.5	8.9	1830121	4	US-09-557-884-1	Sequence 1, Appli
37	202.5	8.9	1830121	4	US-09-643-990A-1	Sequence 1, Appli
38	201.5	8.8	2509	4	US-09-949-016-5225	Sequence 5225, Ap
39	195	8.5	580073	4	US-08-545-528B-1	Sequence 1, Appli
40	182	8.0	924	4	US-09-710-279-2121	Sequence 2121, Ap
41	181	7.9	1230025	4	US-09-198-452A-1	Sequence 1, Appli
42	181	7.9	1230230	4	US-09-438-185A-1	Sequence 1, Appli
43	179	7.8	1230230	4	US-09-438-185A-1	Sequence 1, Appli
44	176.5	7.7	1255	3	US-08-714-918-105	Sequence 105, App
45	176.5	7.7	2255	3	US-09-265-315-105	Sequence 105, App

ALIGNMENTS

RESULT 1
US-09-934-901-15
; Sequence 15, Application US/09934901
; Patent No. 6555353
; PATENT INFORMATION:
; APPLICANT: Koffas, J. Martin
; APPLICANT: Odom, J. Martin
; APPLICANT: No. 6555353ton, Kelley C.
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229, 906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: METHYLOMONAS SP.
US-09-934-901-15

Alignment Scores:

Pred. No.: 7.48e-264 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-701-200-6 (1-437) x US-09-934-901-15 (1-1311)

QY 1 ApgvaIvalThrTrpProGyrHisLeuThrAlaAspIleArgPheCysHisItrPhePhe 20
DB 1 GATGTGTCACATGCGCTTATCATTACCGGCTGATATTCATTGTGTCATTTGTTT 60

QY 21 LeuAenPhaenPhetYrThrleuMetAenlybProlyblyValAlaileuThraJa 40
 DB 61 CTTAACTTAACTTACACGCTCATGAACAACTTAAAGATTGCAATCTGACAGCA 120
 QY 41 GlyGlyLeuAlaProCybLeuAenSerAlaileGlySerleuileGlybArgYrThrglu 60
 DB 121 GGCGGCTTGCGCTGTTGAATTCGCAATCGGATGATTGATCGAACGTTATACCGAA 180
 QY 61 lAaspProserileGlyleileCybYrArgGlyYrYrlybGlyleuLeuengly 80
 DB 181 ATCGATCCATAGCATGAATCATTTGCTATCGGGCGGTTATTAAGCCCTGTTGCTGGGC 240
 QY 81 AaspSerYrProValThrAlaGlyValArglyblybAlaileuGlnArgpheGly 100
 DB 241 GATTCTTAACTCAAGTAAACGCGCAAGCGCTTAAAGCGGCTTCTGCAAGCTTTTGGC 300
 QY 101 GlySerValAlaileGlyAenSerArgValLybLeuThraAenValLybApcYValLybArg 120
 DB 301 GATTCTGATCGGCAACAGCGCGTCAATGACCAATGTCAAAGATCGCGTGAACCGC 360
 QY 121 GlyLeuValLybGlyGlyGlybArgProGlnlybValAlaAapGlnleuValLybAap 140
 DB 361 GGTGGTGTCAAAAGAGGTGAAGATCCGCAAAAAGTCGGCGTGAATTTGGTTAAGAT 420
 QY 141 GlyValAapIleleuHisThrileGlyGlyAapAapThrAenThrAlaAlaAapleu 160
 DB 421 GGTTGTCATATTCGACACATCGCGCGCGATGATACCAATACGACACAGCGGATTTG 480
 QY 161 AlaAlaPheleuAlaArgAenAenYrGlyleuThraValAlaileGlyleuProlySerThrVal 180
 DB 481 GAGGATTCCTGCGCCGAAATATTAAGGATGACCGTATGTTGTTACTTAAACCGTC 540
 QY 181 AapAenAapValPheProIlelybGlnSerleuGlyAlaTrpThrAlaAlaGlyGlnGly 200
 DB 541 GATTAACGATGATTTCCGATCAAGCAATCATAGTGCTTGAATCGCGCGGAGCAAGC 600
 QY 201 AlaArgYrPheMetAenValAlaAlaGlybAenAenAapProArgMetleuileVal 220
 DB 601 GCGCGTTATTTATCATGAACGTGGTGGCGCAAAACAGCCCAACCGCATGCTGATCGTA 660
 QY 221 HisGlyValMetGlyArgAenCybGlyTrpLeuThraAlaileThraAlaGlnGlyYrArg 240
 DB 661 CACGAATGATAGGCGCGTAACTGCGCTGCTGACCGCTGCAACCGCGAGGAATATCGC 720
 QY 241 LybLeuLeuAapArgAlaGlyTrpLeuProGlnleuGlyleuThraArgGlnSerYrGlu 260
 DB 721 AATTAATCTGACCGTGCAGTGGTGGCGGAATGGGTTGATCTGTGAATCTTAATGA 780
 QY 261 ValHisAlaValPheValProGlnMetAlaileAapleuGlnAlaGlybArgleu 280
 DB 781 GTGACGCGGTATTCGTTCCGGAATGGCGATCGAATCGGAAGCCCAACCGCAAGCCCTG 840
 QY 281 ArgGlyValMetAapLybValAapCybValAenIlePheValSerGlnGlyAlaGlyVal 300
 DB 841 CCGGAATGATAGCAAAAGTCATGCGTCAACATTTGTTCCGAAGTGC CGGCGTC 900
 QY 301 GlnAlaileValAlaGlyMetGlnAlaLybGlyGlnGlyValProArgAapAlaPheGly 320
 DB 901 GAAGCATGCTCGCGAAATGACGCGCAAAAGCGAAGAGTCCGCGCATCGCTTGGC 960
 QY 321 HisIleLybLeuAapAlaValAapProGlyYrTrpPheGlyGlyGlnPheAlaGlnMet 340
 DB 961 CACATCAAACTGATCGCGTCAACCTCGTAAATGTTGCGGAGCAAAATTCGCGCAGAG 1020
 QY 341 lAeglyAlaGlyblybThrleuValGlnlybSerGlyYrPheAlaArgAlaSerAlaSer 360
 DB 1021 ATAGGCGCGAAACAACTCGGTAACAAATCGGGATCTTGGCCGCTCTTCTGCTTCC 1080
 QY 361 AenValAapAapMetArgleuIlelybSerCybAlaAapleuAlaValGlyCybAlaPhe 380
 DB 1081 AACGTTGACGACATGCGTTGATCAAAATCGTGGCGCGCATTTGGCGGTCAAGTGGCGGTTCC 1140

QY 361 ArgArgGlySerGlyValAlaileGlyHisAapGlybAapAenGlybAenValleuArgAlaile 400
 DB 1141 CGCCCGCAGTCTGGCGTATCGGTCAACGAAAGCAACCGCAAGCTGTTGGTGGATC 1200
 QY 401 GluPheProArgIlelybGlyGlybArgProPheAenIleAapThrAapTrpPheAenSer 420
 DB 1201 GAGTTTCCGCGCATAGAGGCGGCAACCGTTCAATATCGACACCGATCGTTCAATAC 1260
 QY 421 MetLeuSerGlyIleGlyGlnProlybGlyGlyblybValAlaGlyValSerHis 437
 DB 1261 ATGTTAGACGAATATCGCCAGCCTTAAAGCGGTTAAAGTCGAAGTCAAGCCAC 1311
 RESULT 2
 US-09-934-868-5
 / Sequence 5, Application US/09934868
 / Patent No. 6689601
 / GENERAL INFORMATION:
 / APPLICANT: Koffas, Mattheos
 / APPLICANT: Odem, James M
 / APPLICANT: Schenzle, Andreas J
 / TITLE OF INVENTION: IDENTIFYING METHANOTROPHIC BACTERIAL STRAIN
 / FILE REFERENCE: CL1596 US NA
 / CURRENT APPLICATION NUMBER: US/09/934,868
 / PRIOR APPLICATION NUMBER: 60/229,858
 / PRIOR FILING DATE: 2000-09-01
 / NUMBER OF SEQ ID NOS: 81
 / SOFTWARE: Microsoft Office 97
 / SEQ ID NO 5
 / LENGTH: 1311
 / TYPE: DNA
 / ORGANISM: METHYLOMONAS SP.
 US-09-934-868-5
 Alignment Scores:
 Pred. No.: 7,486-264 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-10-701-200-6 (1-437) x US-09-934-868-5 (1-1311)
 QY 1 AapValAlaThrTrpProYrHisleuThraAapIleArgPheCybHisTrpPhePhe 20
 DB 1 GATGGTCACATGCGCCCTATCATCTTAACGGCTGATTCGATTTGTCATGTTT 60
 QY 21 LeuAenPhaenPhetYrThrleuMetAenlybProlyblybValAlaileuThraJa 40
 DB 61 CTTAACTTAACTTACACGCTCATGAACAACTTAAAGATTGCAATCTGACAGCA 120
 QY 41 GlyGlyLeuAlaProCybLeuAenSerAlaileGlySerleuileGlybArgYrThrglu 60
 DB 121 GGCGGCTTGCGCTGTTGAATTCGCAATCGGATGATTGATCGAACGTTATACCGAA 180
 QY 61 lAaspProserileGlyleileCybYrArgGlyYrYrlybGlyleuLeuengly 80
 DB 181 ATCGATCCATAGCATGAATCATTTGCTATCGGGCGGTTATTAAGCCCTGTTGCTGGGC 240
 QY 81 AapSerYrProValThrAlaGlyValArglyblybAlaileuGlnArgpheGly 100
 DB 241 GATTCTTAACTCAAGTAAACGCGCTTAAAGCGGCTTCTGCAAGCTTTTGGC 300
 QY 101 GlySerValAlaileGlyAenSerArgValLybLeuThraAenValLybApcYValLybArg 120
 DB 301 GATTCTGATCGGCAACAGCGCGTCAATGACCAATGTCAAAGATCGCGTGAACCGC 360
 QY 121 GlyLeuValLybGlyGlyGlybArgProGlnlybValAlaAapGlnleuValLybAap 140
 DB 361 GGTGGTGTCAAAAGAGGTGAAGATCCGCAAAAAGTCGGCGTGAATTTGGTTAAGAT 420
 QY 141 GlyValAapIleleuHisThrileGlyGlyAapAapThrAenThrAlaAlaAapleu 160

[illegible]

```

? PRIOR FILING DATE: 2001-08-22
? PRIOR APPLICATION NUMBER: 60/229,906
? PRIOR FILING DATE: September 1, 2000
? NUMBER OF SEQ. ID NOS: 20
? SOFTWARE: Microsoft Office 97
? SEQ. ID NO. 15
? LENGTH: 1311
? TYPE: DNA
? ORGANISM: METHYLOMONAS SP.
US-10-321-210-15

Alignment Scores:
Pred. No.: 7,48e-264 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
GB: 4 Gaps: 0

```

US-10-701-200-6 (1-437)	x	US-10-321-210-15 (1-1311)	
Qy	1	AspValValThrTyrProTyrHisLeuThrAlaAspIleArgPheCysHisTyrPhePhe	20
Ds	1	GATGTGTCACATGGCGCCTATCATCTTAACGGCTGATATTGCATTTTGCATTGGATT	60
Qy	21	LeuAsnPheAsnPheCysTyrThrLeuMetAsnIleProIleValIleAlaIleLeuThrAla	40
Ds	61	CTTAACCTTAACTCTTACACGCTCAVGAACAACCTTAAAAAGTTGCAATTCACACGCA	120
Qy	41	GIYGIYLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGlu	60
Ds	121	GGCGGCTTGGCGCCTTTGTAATTCGCCAATCGGTACTTGCATCGAACCTGATACCGAA	180
Qy	61	IleAspProSerIleGluIleIleCysTyrArgGlyGlyTyrIleGlyLeuLeuLeuGly	80
Ds	181	ATGCATCTCAGCATGGAATAATCTTTGCTATGGCGCGGTTTAAAGCCGTGTGCTGGC	240
Qy	81	AspSerTyrProValThrAlaGluValArgIleValAlaGlyIleLeuGlnArgPheGly	100
Ds	241	GATTTCTTATCCAGTAAACGGCCGAAGTGGCTAAAAAGGGGGGTGTTCTGCAAACGTTTGGC	300
Qy	101	GlySerValIleGlyValAsnSerArgValIleLeuThrAsnValIleAspCysValIleArg	120
Ds	301	GGTTCTGTATCGGGCAACGCGCTCAAAATTGCACATGTCAAAGACCTCGTGAACGC	360
Qy	121	GlyLeuValIleGlyGlyGlyAspProGlnIleValIleAlaAspGlnLeuValIleAsp	140
Ds	361	GGTTTGCTCAAAAGGGGTGAAGATCCGAAAAAGCTGGCGCTGATCAATTGGTTAAAGAT	420
Qy	141	GlyValAspIleLeuHisThrIleGlyGlyValAspAspThrAsnThrAlaAlaAspLeu	160
Ds	421	GGTGTGATTTCTGCACACATCGCGCGGCATGATACCAATACGCAACGCGCAATTG	480
Qy	161	AlaAlaPheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProIleThrVal	180
Ds	481	GCAGCATTCCTGGCCGAAGAAATTAACGGACTGACCGTCATGGTTTACTTAAACCGTC	540
Qy	181	AspAsnAspValPheProIleIleGlyGlnSerLeuGlyAlaTyrThrAlaAlaGlyGlnGly	200
Ds	541	GATTAACGATATTCCTCCGATCAACCAATCACTAGGTGCTTGACCTGCCCGCAGCAAGGC	600
Qy	201	AlaArgTyrPheMetAsnValValAlaGluAsnAsnAlaAsnProArgMetLeuIleVal	220
Ds	601	GGCGCTTATTTTATGACGATGCGTGGCCGAAAAACAACGCAACCAACCAATGCTGATCTGA	660
Qy	221	HisGluValMetGlyArgAsnCysGlyTyrPheLeuThrAlaAlaThrAlaGlnGluTyrArg	240
Ds	661	CACGAAGTATGGCGCGTAACTGCGGCTGACCGCTGCACCGGCGCAGGAATATCGC	720
Qy	241	IleLeuLeuAspArgAlaGlyIlePheLeuProGlyLeuGlyLeuThrArgGlySerTyrGlu	260
Ds	721	AAATTACTGACCGTCCGACAGTGGTGGCCGGAATTTGGGTTTGATCTGTGAATCTTATGAA	780

```

QY 261 VALHISALAVAlPheValProGluMetAlaIleAspLeuGluAgluAlaIysArgLeu 280
DB 781 GGGACGGCGGATATGCTTCGGAATAAGCGATCGACTTGGAAGCCGAAGCCGCTG 840
QY 281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluGlyVal 300
DB 841 CGCGAAGTATGATGACAAAGTCGATTCGTCGCAATCTTCTTCGGAAGTCCGGCGTC 900
QY 301 GluAlaIleValAlaGluMetGlnAlaIysGlyGlnGluValProArgAspAlaPheGly 320
DB 901 GAAGCTATCGTCGCGGAATGCAAGCCGCAAGCCGCAAGTCCGCGCATCGTTCCGCG 960
QY 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
DB 961 CACATCAAACTGATGAGCGGTCAACCTGCTAAATGTTGCGAGCAATTCGCGCAGATG 1020
QY 341 IleGlyAlaGluLysTrpLeuValGlnLysSerGlyTrpPheAlaArgAlaSerAlaSer 360
DB 1021 ATAGCGCGGAAAAAACCCTGGTACAAAATCGGATCTTCCCGCTCTTCTGCTTC 1080
QY 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
DB 1081 AACGTTGACGACATGCGTTGATCAAAATGTCGCGCATTCGCGGTGCGAGTCCGCTTC 1140
QY 381 ArgArgGluSerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle 400
DB 1141 CCGCGGAGTCTGGCGCTGATCGGTCAAGCAAGAACGCGCAACGTTGTCGTCGATC 1200
QY 401 GluPheProArgIleLysGlyLysProPheAsnIleAspThrAspTrpPheAsnSer 420
DB 1201 GAGTTTCGCGCATCAAGGCGCGCAACCGTTCAATATGACACCGATCGTTCAATAGC 1260
QY 421 MetLeuSerGluIleGlyGlnProLysGlyGlyLysValGluValSerHis 437
DB 1261 ATGTTGACGGAATCGCGCAGCTAAAGCGGTTAAAGTGAAGTCAAGCCAC 1311

RESULT 4
US-10-320-874-15
/ Sequence 15, Application US/10320874
/ Patent No. 6773905
/ GENERAL INFORMATION:
/ APPLICANT: Kofeas, Matheos
/ APPLICANT: Odom, J. Martin
/ APPLICANT: No. 6773905ston, Kelley C.
/ APPLICANT: Ye, Rick
/ TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
/ FILE REFERENCE: CL1619 US NA
/ CURRENT APPLICATION NUMBER: US/10/320,874
/ CURRENT FILING DATE: 2002-12-17
/ PRIOR APPLICATION NUMBER: US/09/934,901
/ PRIOR FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: 60/229,906
/ PRIOR FILING DATE: September 1, 2000
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 15
/ LENGTH: 1311
/ TYPE: DNA
/ ORGANISM: METHYLOMONAS SP.
US-10-320-874-15

Alignment Scores:
Pred. No.: 7,48e-264 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```

```

DB 1 GATGTCACATGCGCCCTATCACTTAACGGCTGATATTCGATTTTGTGATGTTT 60
QY 21 LeuAsnPheAsnPheTrpThrLeuMetAsnLysProLysValAlaIleLeuThrAla 40
DB 61 CTTAACTTAACTTCTTACACCTCTCATGAACAACTTAATAAGTTGCAATCTGACAGA 120
QY 41 GlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTrpThrGlu 60
DB 121 GGCGGCTGGGCGCTGTTGTAATTCGCAATCGGATGATGATGAAACGTTATACCGNA 180
QY 61 IleAspProSerIleGluIleIleCysTrpArgGlyGlyTrpLysGlyLeuLeuLeuGly 80
DB 181 ATGCATCTTACATAGAAATCATTTGCTATCGCGCGGTATTAAGGCTGTTGCTGGGC 240
QY 81 AspSerTrpProValThrAlaGluValArgLysValaGlyValaLeuGlnArgPheGly 100
DB 241 GATTTTATCCAGTAAACCGCGAAGTGGTAAAGGCGGGGTCTTCCACAGTTTGGC 300
QY 101 GlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
DB 301 GGTTCGTGATCGGCACAGCCGCGTCAAAATTGACCAATGTCAAAGACTGCGTGAACGC 360
QY 361 GGTTCGTCAAGAGGGTGAAGATCCGCAAAAAGTCCGCGTGAATCATTTGGTTAAGGAT 420
QY 421 GlyValaAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAlaAlaAspLeu 160
DB 421 GGTGTCGATATTCGACACCATCGCGCGCATGATCAATATACCGCGACGCGGATTTG 480
QY 461 AlaAlaPheLeuAlaArgAsnAsnTrpGlyLeuThrValIleGlyLeuProLysThrVal 180
DB 481 GCAGATTCCTGGCCCAAGAAATTAATTAACGACTGACCGTATGGTTACCTTAAACCGTC 540
QY 181 AspAsnAspValPheProIleLysGlnSerLeuGlyAlaTrpThrAlaAlaGluGlnGly 200
DB 541 GATAACGACGATTTTCGATCAAGCAATCACTAGGTCTTGGACTGCGCGCAGCAAGGC 600
QY 201 AlaArgTrpPheMetAsnValAlaAlaGluAsnAsnAlaAsnProArgMetLeuIleVal 220
DB 601 GCGCGTTATTTTCAATGACAGGTGTCGCAAAAACAACGCCAACCCGATGCTGATCTGA 660
QY 221 HisGluValMetGlyArgAsnCysGlyTrpLeuThrAlaAlaThrAlaGlnGluTrpArg 240
DB 661 CACGAAGTATGAGGCGGTAATCTGCGCTGCTGACCGCTGCAACCGCGCAGAAATATGC 720
QY 241 LysLeuLeuAspArgAlaGluTrpLeuProGluLeuGlyLeuThrArgLysSerTrpGlu 260
DB 721 AATTTACTGACCGCGTCCGAGTGGTCCGGAATTTGGTTTGACTCGTGAATCTTATGA 780
QY 261 ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluAlaLysArgLeu 280
DB 781 GTGCACGCGGTATTCGTTCCGGAATGCGATCGACTCGGAAGCCGAAGCCGAAGCGCTG 840
QY 281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluGlyValaGlyVal 300
DB 841 CGCGAAGTATGAGAAAGTCGATTCGTCCAACATCTTCTTCCGAAGGTCCGCGCTC 900
QY 301 GluAlaIleValAlaGluMetGlnAlaIysGlyGlnGluValProArgAspAlaPheGly 320
DB 901 GAAGCTATCGTCGCGGAATGCAAGCCGCAAGCCGCAAGTCCGCGCATCGTTCCGCG 960
QY 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
DB 961 CACATCAAACTGATGAGCGGTCAACCTGCTAAATGTTGCGAGCAATTCGCGCAGATG 1020
QY 341 IleGlyAlaGluLysTrpLeuValGlnLysSerGlyTrpPheAlaArgAlaSerAlaSer 360
DB 1021 ATAGCGCGGAAAAAACCCTGGTACAAAATCGGATCTTCCCGCTCTTCTGCTTC 1080
QY 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
DB 1081 AACGTTGACGACATGCGTTGATCAAAATGTCGCGCATTCGCGGTGCGAGTCCGCTTC 1140

```

QY 381 ArgArgGluSerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle 400
 DB 1141 CGCCGGAGCTCTGGCTGATCGCTCAGCAGAGCAACGCAAGCTGTTGGCTGGATC 1200
 QY 401 GlnPheProArgIleGlySerGlyLysProPheAsnIleAspThrAspTrpPheAsnSer 420
 DB 1201 GAGTTTCCCGCGCATCAAGGCGGCAACCGTTCAATATCGACACCACTGGTTCAATAGC 1260
 QY 421 MetLeuSerGlnIleGlyGlnProLysGlyLysValGluValSerHis 437
 DB 1261 ATGTTGAGGCAATCGCGCCAGCTTAAGGCGGTAAAGTCGAGTCAGCCAC 1311
 RESULT 5
 US-08-416-870C-3
 ; Sequence 3, Application US/08416870C
 ; Patent No. 5834862
 ; GENERAL INFORMATION:
 ; APPLICANT: HIYOSHI, TORU
 ; APPLICANT: MINE, TOSHIKI
 ; APPLICANT: KASAKURA, KEISUKE
 ; APPLICANT: TYSON, ROBERT HUM
 ; APPLICANT: PAGE, ANTHONY MILES JOHN
 ; TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
 ; TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTO
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 ; STREET: PO BOX 747
 ; CITY: FALL CHURCH
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22040-0747
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/416,870C
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURPHY JR, GERALD M
 ; REGISTRATION NUMBER: 28,977
 ; REFERENCE/DOCKET NUMBER: 760-195P (PCT)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 205-8000
 ; TELEFAX: (703) 205-8050
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1778 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 19..1467
 ; US-08-416-870C-3
 Alignment Scores:
 Pred. No.: 5,29e-22 Length: 1778
 Score: 270.00 Matches: 113
 Percent Similarity: 40.59% Conservative: 53
 Best Local Similarity: 27.63% Mismatches: 175
 Query Match: 11.82% Indels: 68
 DB: 1 Gaps: 15
 US-10-701-200-6 (1-437) x US-08-416-870C-3 (1-1778)
 QY 36 AlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIle 55

DB 301 GCTATAGTAAACATCTGTGGTTTATGCTCTGGCGCTAAACACAGTATCAAGGAAATGTT 360
 QY 56 GlnArgTyThrGlnIleAspProSerIleGlnIleIleCysTyArgGlyValLys 75
 DB 361 TGGCAGCTTATTCACATGATGTGTGTACCAAGTTCTTGGAGTTGATGAGGTCACAG 420
 QY 76 GlnLeuLeuGlyAspSerTyProValThrAlaGluValArgLysAlaGlyVal 95
 DB 421 GGTTTTACTCAAAAACACATCACTTGACTCCAAAGTT-----GGAATGAC 471
 QY 421 GGTTTTACTCAAAAACACATCACTTGACTCCAAAGTT-----GGAATGAC 471
 DB 96 LeuGlnArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThrAsnValLys 115
 DB 472 ATCCATTAACGTGTGTGTACCAATTATGGCACCTCTCGT-----510
 QY 116 AspCysValLysArgGlyLeuValLysGlyGlyGlnAspProGlnLysValAlaAsp 135
 DB 511 -----GGGGCGCATGATTAACCAAAAGATA--GTTGAC 540
 QY 136 GlnLeuValLysAspGlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThr 155
 DB 541 AGATTTCAGGATCGTGTATTCATCACTGTTTATATGAGAGACCGGTACTCAAAAG 600
 QY 156 AlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTyrglyLeuThrValIleGly 175
 DB 601 GGAGCAGCTGTATTATATCAGAGAGTGAAGAGCGGCTTAACCTGATGAGCTGGG 660
 QY 176 LeuProLysThrValAlaAspAsnAspValPheProIleLysGlnSerLeuGlyValATPTTr 195
 DB 661 ATTCCTTAAGACATGATATGATATGACATTCGTCATGATGATGATGTTGTTTGAACAG 720
 QY 196 AlaAlaGlnGlnGlyAlaArgTyThrPheMetAsnValAlaGlnAsnAsnAlaAsnPro 215
 DB 721 GCTGTGAAGAGCTCAACGTGCTATTAATGCTGCACATGTGAGGCTGAAGCTGAG 780
 QY 216 ArgMetLeuIleValHisGlnValMetGlyArgAsnCysGlyTyThrLeuThrAlaIle 235
 DB 781 AATGCATAGGGGTGTGTCAACTATGAGCAGCTATGAGATTCATCCCA-----831
 QY 236 AlaGlnGlnTyArgLysLeuLeuAspArgAlaGluTrpLeuProGlnLeuGlyLeuThr 255
 DB 832 -----ATGATGCAACT 843
 QY 256 ArgGluSerTyrglyValHisAlaValPheValProGluMetAlaIleAspLeuGluAla 275
 DB 844 TTGGCTAGTCGAGATGTTATTTATGTTTAACTCTGAAATCACTTTTATCTTGAGGA 903
 QY 276 GluAla-----LysArgLeuArgGluValMetAspLysValAsp 288
 DB 904 GAAAGTCGACTTTTAGAATATGTAGAAAACGCTCAAGACAGATGACATGCTC---960
 QY 289 CysValAsnIlePheValSerGlyGlyAlaGlyValGluAlaIleValAlaGluMetGln 308
 DB 961 -----ATCGTTGTACAGAGGTGTGTGTGAGACCTGCTTCACAGAAAACCTTG 1011
 QY 309 AlaLysGlyGlnGluValProArgAspAlaPheGlyHisIleLysLeuAspAlaValAsn 328
 DB 1012 AAA-----ACTTCAACCGCAAAAGATGCTTCTGGAATAAATCTACTCAAGATGTC---1062
 QY 329 ProGlyLysTrpPheGlyGluGln-----PheAlaGlnMetIleGlyAlaGlu 344
 DB 1063 ---GATTTGTGATTTCTGATAAGATTAGGCTCACTTTGCTAAATTCCTCCCATGCGCT 1119
 QY 345 LysThrLeuValGln---LysSerGlyTyThrAlaArgAlaSerAlaSerAsnValAsp 363
 DB 1120 ATTACTCTCAATAATACATAGATCCAACTTACAGATCCGCGGTTCACAGTATGATCAT 1179
 QY 364 AspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArgArgGlu 383
 DB 1180 GATTAAT-----GTATATGCACTCTCTCTGCAAAAGTTGTTCATGAGATGATG 1230
 QY 384 SerGlyValIleGlyHisAspGluAsp-----AsnGlyAsnValLeuArgAlaIleGlu 401

Db 1231 GGGGGCTACACCGGCTTCAACAAGTGGCTTGTCAATGTAGA---CAGACTTATATCCA 1287
 Qy 402 PheProArgIleuysGlyIysPheProheanIleaspThrAsp-----TrpPhean 419
 Db 1288 TTTAATCGTATCATCGAAGACGATATACGTTGTATACCGATGATGATGGCAAG 1347
 Qy 420 SerMetLeuSerGluIleGlyInPro 428
 Db 1348 ---CTTGTGTATCCACCAACCAACCA 1371

RESULT 6

US-08-416-870C-5
 / Sequence 5, Application US/08416870C
 / Patent No. 5824862

GENERAL INFORMATION:
 APPLICANT: HIYOSHI, TORU
 APPLICANT: MINE, TOSHIKI
 APPLICANT: KASAKA, KEISUKE
 APPLICANT: TYSON, ROBERT HEN
 APPLICANT: PAGE, ANTHON MILLS JOHN
 TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
 TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTO
 NUMBER OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDE
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 STREET: PO BOX 747
 CITY: FALL CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22040-0747

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/416,870C
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURPHY JR, GERALD M
 REGISTRATION NUMBER: 28,977
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1624 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 3..1409
 US-08-416-870C-5

Alignment Scores:
 Pred. No.: 2.09e-21 Length: 1624
 Score: 264.50 Matches: 109
 Percent Similarity: 40.54% Conservative: 56
 Best Local Similarity: 26.78% Mismatches: 173
 Query Match: 11.58% Indels: 69
 DB: 1 Gaps: 16

US-10-701-200-6 (1-437) x US-08-416-870C-5 (1-1624)

Qy 37 ILeuLeuThrAlaGlyIleuAlaProCybLeuAsnSerAlaIleGlySerLeuIleGlu 56
 Db 222 ATTGTCAATGTCAGAGACTGTGCTCGACTGACACTGATCACTGTCTATTAGGAAATTGTTGT 281

Qy 57 ArgIYrThrGluIleAspProSerIleGluIleIleCysTYrARgGlyGlyTYrIleGly 76
 Db 282 GGCCTAAATGACATGATATGTCATGATGAGGACTGGAATTCAGGGTGGATATGAGCT 341
 Qy 77 LeuLeuLeuGlyAspSerTYrProValThrAlaGluValArgYsLysAlaGlyValIleu 96
 Db 342 TTCTATGCTTGTAAACACCATGACTGACTGATCA-----AACAGTAAACAGCAAT 392
 Qy 97 GluArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThrAsnValIleAsp 116
 Db 393 CACAAAAGGGGTGAACCTGTTCTGGGACATCAGCT----- 428
 Qy 117 CysValLysArgGlyLeuValLysGluIleGluAspProGlnLysValAlaIleAspGln 136
 Db 429 -----CGAGCCATGACACCATGACATG---CTTACAC 461
 Qy 137 LeuValLysAspGlyValAspIleLeuIleThrIleGlyIleAspThrAsnThrAla 156
 Db 462 ATCCAGATCGGTGATTAATCAAGTTATGTAATGTTGATGATGATCAAAAGGGGT 521
 Qy 157 AlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTYrGlyLeuThrValIleGlyLeu 176
 Db 522 GCAGAGATGATTTTGAAGATTTGAAGACGTGCTCAAGGTGCTGCTGCTGCTGCAAT 581
 Qy 177 ProIYrThrValAspAsnAspValPheProIleLysGlnSerLeuGlyAlaIleThrAla 196
 Db 582 CCNAAAGACGATTAAGATTAATGACATGATTAATGACATGATTTGTTTGCACACTGA 641
 Qy 197 AlaGluGlnGlyAlaArgTYrPheMetAsnValAlaGluAsnAsnAlaAsnProArg 216
 Db 642 GTTGAAGAGCCCAACGTGCAATTAATGCTGCTCATGTGAAGCTGAAGCCCGAGAT 701
 Qy 217 MetLeuIleValHisGluValMetGlyValArgAsnCYsGlyTYrPleuThrAlaIleAla 236
 Db 702 GGTATAGGCTCGTAAAGCTAATGATGATGATGATGATGATGATGATGATGATGATGAT 749
 Qy 237 GlnGluTYrArgGlyLeuLeuAspArgAlaGluTYrPleuProGluLeuGlyLeuThrArg 256
 Db 750 ---CACTATGCTACTTACCCAGCAGA----- 773
 Qy 257 GluSerTYrGluValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGlu 276
 Db 774 -----GACGTGATGTTGTTGTTGATTCAGAGTCACTTCTATGTGAAGCTGA 824
 Qy 277 Ala-----LysArgLeuArgGluValMetAspLysValAspCys 289
 Db 825 GGTGCTTTTATGATTTTGAAGAGCTGCAAGAGAAATGTCATATGCTT----- 878
 Qy 290 ValAsnIlePheValSerGluGlyAlaGlyValGluAlaIleValAlaGluMetGlnAla 309
 Db 879 -----ATCGTTGTTGCGAGGGTCCAGGCGCAAGAACTTATTATGAACAAAGATCA 932
 Qy 310 LysGlyGlnGluValProArgAspAlaPheGlyHis-----IleLysLeuAspAlaValAsn 328
 Db 933 ATGGGG-----AAAGATCTTCAGCAATTCGATTCCTTATGATGTT----- 974
 Qy 329 ProGlyLysTYrPheGlyGlnGluPheAlaGluMetIleGlyAlaGluTYrThrLeuVal 348
 Db 975 ---GGTCTTGTGTTATCTCAAAAGATAAAGAGCATTTCAAGAAATCAAGACTATATA 1031
 Qy 349 Gln-----LysSerGlyTYrPheAlaArgAlaSerAlaSerAsnValAspAsp 364
 Db 1032 AATCTCAAGTATATGATCTTACATATCATGATATGATCTTCTTCAATCATCTGAC 1091
 Qy 365 MetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArgArgLysSer 384
 Db 1092 AATGTGATTCACACTGTGTCACACAGGCTGTTTCATGAGCCATGCTGATCACT 1151
 Qy 385 Gly-----ValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIleGluPhePro 403
 Db 1152 GGTTCACGCTTGGC-----CAAGTAAATGTCGAGCATGTC---TATATCCGTTTAC 1202
 Qy 404 ArgIle-----LysGlyGlyLysPheProheanIleAspThrAspTrpPheAsnSerMet 421


```

Db      1203 AGGATTCACAGAGAAGCAGAAACAATTCTTCAATTACTGATGAGGTGGCCAGAGA---CTT 1259
QY      422 LeusergluileglyInPro 428
        |||||
        |||||
Db      1260 CTCCTCACCAACCACCA 1280
        |||||

RESULT 7
US-08-416-870C-1
; Sequence 1, Application US/08416870C
; Patent No. 5824862
; GENERAL INFORMATION:
APPLICANT: HIYOSHI, TORU
APPLICANT: MINE, TOSHITSUGU
APPLICANT: KASAOKA, KEISUKE
APPLICANT: TYSON, ROBERT HWM
APPLICANT: PAGE, ANTHONY MILES JOHN
TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
TITLE OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDER STRESS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
STREET: PO BOX 747
CITY: PALM CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,870C
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 760-195P(PCT)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1978 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURES:
NAME/KEY: CDS
LOCATION: 133..1587
US-08-416-870C-1

Alignment Scores:
Pred. No.:          Length:    1978
Score:              Matches:     113
Percent Similarity: 38.41%
Best Local Similarity: 25.68%
                    Mismatches:   179
Query Match:       Indels:       92
DB:                Gaps:         17

US-10-701-200-6 (1-437) x US-08-416-870C-1 (1-1978)
QY      7 TyrHisLeuThrAlaAspIleArgPheCysHisTrpPhePheLeuAsnPheAsnPheTyr 26
        |||||
        |||||
        ..:::
Db      388 TATTTCAGTTGGAGTGATGTCGTCTGT----- 417
        |||||
QY      27 ThrIleMetAlanIlyeProIlybysValAlaIleLeuThrAlaGlyIleuAlaProCys 46
        |||||
        |||||
        -ATTGTAACCTGTGGTGGTTGGTGGCCCTGG 447
        |||||

```

47 LeuAenSerAlaIleGlySerLeuIleGluArgTrpThrgluIleAerProSerIleGlu 66
448 CTTAAACACAGAGATCAAGAGATGTTGTACATAGACCTCGATTATATAGTATGAGATCCAA 507
67 IleIleCysTrpArgGlyGlyTrpGlyGlyLeuIleuSerGlyAerSerTrpProValThr 86
508 GTCTTTGGTATCGATGGAGGCTACAGGGGTTTCTATTCCAAAGATATCATCATTTGACA 567
87 AlaGluValArgGlyValArgValIleuGlnArgPheGlyGlySerValIleGlyAsn 106
568 CCA-----AAGACTGTTAATGACATTCATTAAGCGTGGTGATACAAATCTTGATCA 618
107 SerArgValIlySerLeuThrAenValIlyAerCyseValIlySerGlyLeuValIyGlyGly 126
619 TCACGA-----GGAGGC 630
127 GluAerProGlnIySerValAlaAlaAerGlnIleuValIyAerGlyValAerIleuHis 146
631 CATATATCCAAAGATTT--GTTGACAGCATACAGACCGTGAAATTAATTCAGCTATAT 687
147 ThrIleGlyIyAerAerThrAenThrAlaAlaAerPheuAlaAlaPheIleuAlaArg 166
688 ATATCCGTGGTGTATGAACTCAGAAAGACGACCTTATATATGAGAAATCAGGCGG 747
167 AenAenTrpGlyLeuThrValIleGlyLeuProIySerThrValAerAerAerValPhePro 186
748 CGTGCTCCAAAGTAATTTGTTGCTGGAGATCCCAAGACAAATTGATATGATATCCCTGTT 807
187 IleIySerGlnSerLeuGlyAlaTrpThrAlaAlaGlnGlnGlyAlaArgTrpPheMetAsn 206
808 ATCAACAAGTCATTTGTTGTATATCGCTGTACAGAGGCTCAACGTCATTAATGCA 867
207 ValValAlaGluAenAenAlaAerProArgMetLeuIleValIleGluValMetGlyArg 226
868 GCTCATGTGTGAAGCTGAAGTCGAAAGTCGAAATGCTATTGGTGTGCGAAGCTAATGGACGC 927
227 AenCyseGlyTrpLeuThrAlaAlaThrAlaGlnIuTrpArgIyLeuLeuAerPArgAla 246
928 TATAGTGATTCATTCGCA----- 945
247 GluTrpLeuAerProGlnIleuGlyLeuThrArgGluSerTrpGluValIleAlaValPheVal 266
946 -----ATGTATGCCACTTTGGCGCAGAGATGTGATCTCTGTTTAAT 990
267 ProGluMetAlaIleAerPheGluAlaGluAla-----LysArg 279
991 CCAAGATCAACCTTTTATCTTTGAAAGAGATGTGGACTCTTTGAAATCATTTGAAAAAAG 1050
280 LeuArgGluValMetAerPlySerValAerCyseValAenIlePheValSerGlnIyAlaGly 299
1051 CTCAAAGAAATATGGGCATCGTT-----ATTGGATATGCCGAAAGGACGAGG 1098
300 ValGluAlaIleValAlaGlu--MetGlnAlaIySerGlyGlnGluValProArgAspAla 318
1099 CAAGAAGCTCTTCTTGCAAGAAAGAAATGCCATGCCATTCTTAAGTATATAGTCCGACT 1149
319 PheGlyHisIleIySerLeuAerPAlaValAerProGlyIySerTrpPheGlyGlnPheAla 338
1150 TCGGGGAACAGCTCTTCAGAGATGT-----GTTTGGATATTTCCCAAAAAATCAGG 1203
339 GlnMetIleGlyAlaGluIySerThrLeuValGln-----LysSerGly 352
1204 GATCATTTT---GCTACAAAAAACTAAGATGCCATTCTTAAGTATATAGTCCGACT 1266
353 TyrPheAlaArgAlaSerAlaSerAenValAerAerPheArgLeuIleIySerCyseAla 372
1261 TACATGATCTCGCTGTCTTCCAAAGTAATGCCCTGTATAT-----GTAATATGCACT 1311
373 AserLeuAlaValGlyCyseAlaPheArgArgGluSerGlyValIleGlyHisAerGluAer 392
1312 CTTCTTGCTCAAAAGTTGTGTTCAATGAGCAATGGCAGGCTACACAGGTTTCACTCAGCA 1371

QY 393 -----AenGlyAsnValLeuArgAlaIleGluPheProArgIleGlyGlyLeuPro 410
DB 1372 CTGTCCAAATGGTCG---CAGACTATATACCATTCATTCATCCAGAAACAAAT 1428
QY 411 PheAnileAspThrAsp-----TTPheAnSerMetLeuSerGluIleGlyGlnPro 428
DB 1429 ATGGTCGTTATATACACAGATGTGGACGCT---CTTCTTTCGCAACCATCAACCA 1485
RESULT 8
US-09-902-540-8974
Sequence 8974, Application US/0902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 8974
LENGTH: 1038
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-8974
Alignment Scores:
Pred. No.: 2,72e-21 Length: 1038
Score: 261.00 Matches: 96
Percent Similarity: 43.90% Conservative: 66
Best Local Similarity: 26.02% Mismatches: 141
Query Match: 11.43% Indels: 66
Gaps: 16
US-10-701-200-6 (1-437) x US-09-902-540-8974 (1-1038)
QY 34 LysValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySer 53
DB 4 AAAGTCGCGCTGCTACCGCGGGGTGATGCTCGCGCTTAACCCCTCATCCGCGC 63
QY 54 LeuIleGluArgTyrThrGluIleAspProSerIleGluIleCysTyrArgGlyGly 73
DB 64 GTCGTCGCGCGCGCAACGCC-----CACGGCTTGAGATGATGGCTCCGAGATGCT 117
QY 74 TyrLysGlyLeuLeuLeuLysAspSerTyrProValThrAlaGluValArgLysAla 93
DB 118 TGGAAAGGGGTGTGTGAGAGACAACTTCGCGCTCACCGGTGAACC-----ACGTCC 171
QY 94 GlyValIleuGlnArgPheGlyGlySerValIleGlyAsnSerArgVal-----Lys 110
DB 172 GGAATCTCTCACCAG---GGCGAACCATCTCTCGGACTCGCGCTCAACCCGTTCAAG 228
QY 111 LeuThrAsnValLysAspCysValLysArgGlyLeuValLysGluGlyLysAspProGln 130
DB 229 GTCGAAAACGGGCTGAGCGCGCTCAAGCGCGC----- 261
QY 131 LysValAlaAlaAspGlnLeuValLysAspGlyValAspIleLeuHisThrIleGlyGly 150
DB 262 -----ATCGAACGCAACGGCGCTCATCCGCGCATTTGCTGCG 303
QY 151 AspAspThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnSerTyrGly 170
DB 304 GAAGGACGCTGTGCGCGCGCGCGCGCATG-----TCGCAAGAAAGA 345
QY 171 LeuThrValIleGlyLeuProLysThrValAspAsnAspValPheProIleLysGlnSer 190
DB 346 CTGGCGCATGTGCTGTCGGAAGACCATCGAACAGATCAACGCCGCGCATTCACAC 405
QY 191 LeuGlyAlaIleTyrThrAlaAlaGluGlnGlyLysArgTyrPhe-----MetAsnValVal 208

DB 406 TTCGCTTCGACACCGCGCGCTGCCCATTCGCCACGAGGCCATTGACCGGCTGCATTCAC 465
QY 209 AlaGluAsnAsnAlaAsnProArgMetLeuIleValIleGluValMetGlyArgAsnGly 228
DB 466 GCGAGTGGAC-----AAGCGCTCATCTGTCGAGGTATGCGCGCTCACG 516
QY 229 GlyTyrLeuThrAlaAlaThrAlaGlnGluTyrArgLysLeuLeuAspArgAlaGluTyr 248
DB 517 GCGTGGATTGGCCAC-----TACGCGGCGCATCTGCGCGCGCGGAC--- 558
QY 249 LeuProGluLeuGlyLeuThrArgLysSerTyrGluValIleAlaValPheValProGlu 268
DB 559 -----GTCACTCTGATGCGGAG 576
QY 269 MetAlaIleAspLeuGluAlaGluAlaLysArgLeu-----ArgGluValMetAspLys 286
DB 577 ATTCGCGGACCTCGCGAAGGTGCGGACGACATCGACCGCGCGCGCGCGCGCG 636
QY 287 ValAspCysValAsnIlePheValSerGluGlyAlaGlyValAlaIleValAlaGlu 306
DB 637 ACCTTCTCCATT---GTGCTGTGCGGAGGCTACGCGCATCAAGCTGTGCGGACG 693
QY 307 MetGlnAlaLysGlyGlnGluValProArgAspAlaPheGlyHisIleLeuLeuAspAla 326
DB 694 CAGAGCAGCTCTGCACACGCGCGCTGACGAGGACGAGCGCGCGCTCGGTGC 753
QY 327 ValAsnProGlyLysTyrPheGlyGluGlnPheAlaGlnMetIleGlyValAlaGluTyr 346
DB 754 GTG-----GGACCATCTGCGCGCGGACGAATCGAGCGCGCGCGCTTCGAG---ACG 804
QY 347 LeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSerAsnValAspMetArg 366
DB 805 CGCGTCTCGCTGCGCGCGCACATCCACGCGCGCGCGCGCGCGCGCGCGCGCG 864
QY 367 Leu-----IleLysSerCysAlaAspLeuAlaValGluCysAlaPheArg 381
DB 865 CTCGCGACCGCTACGCGGCTCACGCTGC---GACATGTGCGCGCGCGCGCGGATTCG 921
QY 382 ArgGluSerGlyValIleGlyHisAsp 390
DB 922 AAGATGCGCGCGCTCGAGGACACAC 948
RESULT 9
US-09-902-540-958/c
Sequence 958, Application US/0902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 958
LENGTH: 10391
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-958
Alignment Scores:
Pred. No.: 1.03e-19 Length: 10391
Score: 261.00 Matches: 96
Percent Similarity: 43.90% Conservative: 66
Best Local Similarity: 26.02% Mismatches: 141
Query Match: 11.43% Indels: 66
Gaps: 16

US-10-701-200-6 (1-437) x US-09-902-540-958 (1-10391)

QY 34 LysValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySer 53
 DB 3602 AAAGTGGCCGCTGTCACCGCGGGGTGACTGCCCGGCTGAACCCGCTATCCGCGGC 3543
 QY 54 LeuIleGluArgTyrThrGluIleAspProSerIleGluIleIleCysTyrArgIleGly 73
 DB 3542 GTTCGTCGCCGCCCAACGCC-----CACGCGCTTGAGATGTGGCCCTCCAGATGTGT 3489
 QY 74 TyrLysGlyLeuLeuLeuGlyAspSerTyrProValThrAlaGluValArgLysLysAla 93
 DB 3488 TGGAGGGGTTGTTGGAGAGACACACTTCGCCCTACCGCTGAAAC-----ACGTCC 3435
 QY 94 GlyValLeuGlnArgPheGlyGlySerValIleGlyAsnSerArgVal-----Lys 110
 DB 3434 GGAATCTCCACCGG---GGCGGAACCATCTCGGCACTCGCGGCTCAACCCGCTCAAG 3378
 QY 111 LeuThrAsnValLysAspCysValLysArgGlyLeuValLysGluGlyLysAspProGln 130
 DB 3377 GTTCGAACAGCGGCTGGAGCGCGCTCAAGCGCGCC----- 3345
 QY 131 LysValAlaAlaAspGlnLeuValLysAspGlyValAspIleuHisThrIleGlyGly 150
 DB 3344 -----ATCGAACCAACGGCATCCACCGCTCATCCGCTATGGTGC 3303
 QY 151 AspAspThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTyrGly 170
 DB 3302 GAAGGACCGCTGTCGCCGCCCGCCACGCCCATG-----TCGCAAGAGAA 3261
 QY 171 LeuThrValIleGlyLeuProLysThrValAspAsnAspValPheProIleLysGlnSer 190
 DB 3260 CTGCGCATGCTGGTGTGCTGCGGAAGACCATCGAACAGACATCAACCCACGACCTTCACC 3201
 QY 191 LeuGlyAlaIleThrAlaAlaGluGlnGlyAlaArgTyrPhe-----MetAsnValAl 208
 DB 3200 TTCCGCTTGAACCGCGCTCCGCCATCGCCACCGAGCCATTGACCGGCTCCACC 3141
 QY 209 AlaGluAsnAsnAlaAsnProArgMetLeuIleValIleGlyValMetGlyArgAsnCys 228
 DB 3140 GGGGAGTCCAC-----AAGCCGTCATGTCGTGGAGAGTGTATGGCCCTCACGTC 3090
 QY 229 GlyTyrPheThrAlaAlaThrAlaGlnGlyTyrArgLysLeuLeuAspArgIleGluTyr 248
 DB 3089 GGCTGATTCACAC-----TACGCGGCGATCGCTGCGCGCGGAC----- 3048
 QY 249 LeuProGluLeuGlyLeuThrArgGluSerTyrGluValAlaAlaValPheValProGlu 268
 DB 3047 -----GTCACTCTGTGGCCGAG 3030
 QY 269 MetAlaIleAspLeuGluAlaGluAlaLysArgLys-----ArgGluValMetAspLys 286
 DB 3029 ATTCCCGCGGACCTCGCGAAGGTGGCGGACATCAAGCGCGCCACCGCGGCGGCGCC 2970
 QY 287 ValAspCysValAsnIlePheValSerGlyGlyAlaGlyValGluAlaIleValAlaGlu 306
 DB 2969 ACCTTCTTCATG---GTGCTGTGGCGGAGGAGTACCGCATCAAGCTGTGCGGCGACAG 2913
 QY 307 MetGlnAlaLysGlyGlnGluValProArgAspAlaPheGlyHisIleLysLeuAspAla 326
 DB 2912 CAGGAGCAGCTGTCACAGCGCGCATCGAGCAGGAGCGCGCGCTCGGTGTC 2853
 QY 327 ValAsnProGlyLysThrPheGlyGluGlnPheAlaGluMetIleGlyAlaGluLysTyr 346
 DB 2852 GTG-----GGACCATCTCTGGCGGACGAATGAGACGGCGCACCGGCTTCGAG---AGC 2802
 QY 347 LeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSerAsnValAspPheMetArg 366
 DB 2801 CCGGTGTCTGTGGCGACATCAACGCGGCGGCGCGCCACCGCGCATGACCGCGTG 2742
 QY 367 Leu-----IleLysSerCysAlaAspLeuAlaValGluCysAlaPheArg 381
 DB 2741 CTCGCCACCCGCTACGCGCTGC---GACATGTGTGCGCGCGGCGAGTTTCGGG 2685

QY 382 ArgGluSerGlyValIleGlyHisAsp 390
 DB 2684 AAGATGGCCGCGCTGCGAGCAACAGC 2658

RESULT 10
 US-08-416-870C-9
 / Sequence 9, Application US/08416870C
 / Patent No. 5824862
 / GENERAL INFORMATION:
 / APPLICANT: HIYOSHI, TORU
 / APPLICANT: KASAKURA, KEISUKE
 / APPLICANT: TYSON, ROBERT HUM
 / APPLICANT: PAGE, ANTHONY MILES JOHN
 / TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
 / TITLE OF INVENTION: 1-PHOSPHORANSEASE ORIGINATING FROM PLANT, RECOMBINANT VECTOR
 / TITLE OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDER
 / NUMBER OF SEQUENCES: 23
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 / STREET: PO BOX 747
 / CITY: FALL CHURCH
 / STATE: VA
 / COUNTRY: USA
 / ZIP: 22040-0747
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / OPERATING SYSTEM: IBM PC compatible
 / SOFTWARE: PatentIn Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/416,870C
 / FILING DATE:
 / CLASSIFICATION: 435
 / ATTORNEY/AGENT INFORMATION:
 / NAME: MORPHY JR, GERALD M
 / REGISTRATION NUMBER: 28,977
 / REFERENCE/DOCKET NUMBER: 760-195P (PCT)
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (703) 205-8000
 / TELEFAX: (703) 205-8050
 / INFORMATION FOR SEQ ID NO: 9:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1558 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: double
 / TOPOLOGY: linear
 / MOLECULE TYPE: cDNA to mRNA
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 1..1278
 / US-08-416-870C-9

Alignment Scores:
 Pred. No.: 6.8e-21 Length: 1558
 Score: 260.00 Matches: 107
 Percent Similarity: 41.15% Conservative: 65
 Best Local Similarity: 25.60% Mismatches: 174
 Query Match: 11.38% Indels: 72
 DB: 1 Gaps: 16

US-10-701-200-6 (1-437) x US-08-416-870C-9 (1-1558)

QY 37 IleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGlu 56
 DB 148 ATTGTACATGTGTGGCTTGTGTCAGGCTTAATACGTATCAGAGAAATCGTTGT 207
 QY 57 ArgTyrThrGluIleAspProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGly 76
 DB 208 GGATTTGCTTACATGATGTGTCAAGAAATCTTGGCATTGAGAGGAGCTTACAGAGC 267
 QY 77 LeuLeuLeuGlyAspSerTyrProValThrAlaGluValArgLysAlaGlyValLeu 96

QY 208 VALAAGLAAASNAALAAENProAArgMcLLeuIleValHicGluValMetGlyArgAsn 227
 Db 1147 GATGCCACCGCTAAATCCACATCCGTCCTTGTTGTT---GAAGTTATGGGTAGACAT 1203
 QY 228 CysGlyTrrPleuThr-----AAlaAlaThrAlaGlnLutYrArgLysLeu 243
 Db 1204 TGTGTGTGTGTGCTTGAAGCTGTATGCTATGCTACCGTCCGATTAAC----- 1251
 QY 244 AepAArgAlaGluTrrPleuProGluLeuGlyLeuThrArgLysSerGlyValHicAla 263
 Db 1251 ----- 1251
 QY 264 ValPheValProGluMetAlaIleAepLeuGluAlaGluAlaLysArgLeuArgGluVal 283
 Db 1252 ATTTTATTCAGAAAGACCTGTCTCTACAGAAATGACAGACAGATTGAAGAGAGTG 1311
 QY 284 MetAepLysValAepCysVal-----AsnIlePheValSerGluGlyAla 298
 Db 1312 TGCCAAAGACACAGAAAGTAGAGTAGAAGAAATTAACACAAATTATGTCCTGAAGGTGCT 1371
 QY 299 -----GlyValGluAlaIleValAlaGluMet 307
 Db 1372 TTAGATGATCAATTAAACCTGTACTGCCAATGACGTCAAGAGTCTTGATTAATG 1431
 QY 308 GlnAlaLysGlyGlnGluValProArgAepAlaPheGlyValIleLysLeu-----Aep 325
 Db 1432 -----GCTCTAGACACACAGAGTAGACCATTCATGTCACGTTCAAGAGGTGTACA 1482
 QY 326 AlaValAepProGlyLysTrrPheGlyGluGlnPheAlaGlnMetIleGlyValGluLys 345
 Db 1483 GCTGTGTCTCATGACAGATGTTAGTACT-----CTCAAGGTGTCTCATGTTAG 1536
 QY 346 ThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSerAsn----- 361
 Db 1537 GCCGTTCTGGA-----TTTACCCTCGAAATCTCTTCTCATTAATGTTAT 1584
 QY 362 -----ValAepAepMetArgLeuIleLysSerCys 371
 Db 1585 TTAGAAACACAGATTAATTGAATGCAATGCTTGTAATCTGTGAAGTTGATTAATCTGTT 1644
 QY 372 AlaAepLeuAlaValGluCysAlaPheArgArgLysSerGlyValIleGlyHicAepGlu 391
 Db 1645 GCC-----ACTGCATTCGAAACAAAGATTTTC 1671
 QY 392 AepAepGlyAepValLeuArgAlaIleGluPhe 402
 Db 1672 GATTAAGCAATTTCTTTAAGACACAGAAATTT 1704
 RESULT 12
 US-09-902-540-788/c
 / Sequence 788, Application US/09902540
 / Patent No. 6833447
 / GENERAL INFORMATION:
 / APPLICANT: Goldman, Barry S.
 / APPLICANT: Hinkle, Gregory J.
 / APPLICANT: Slater, Steven C.
 / APPLICANT: Mlegend, Roger C.
 / TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 / FILE REFERENCE: 38-10(15849)B
 / CURRENT APPLICATION NUMBER: US/09/902,540
 / PRIOR FILING DATE: 2001-07-10
 / PRIOR APPLICATION NUMBER: 60/217,883
 / NUMBER OF SEQ ID NOS: 16825
 / SEQ ID NO 788
 / LENGTH: 6196
 / TYPE: DNA
 / ORGANISM: Myxococcus xanthus
 US-09-902-540-788
 Alignment Scores: 1.73e-17 Length: 6196
 Pred. No.:

Score: 239.50 Matches: 97
 Percent Similarity: 42.41% Conservative: 65
 Best Local Similarity: 25.39% Mismatches: 127
 Query Match: 10.49% Indels: 93
 DB: 4 Gaps: 17
 US-10-701-200-6 (1-437) x US-09-902-540-788 (1-6196)
 QY 6 ProTyrHicLeuThrAlaAepIleAArgPheCysHicTrrPhe-----PheLeuAsnPhe 23
 Db 2269 CCTACATACAGGTCGGA-----CGTTCTGTCAATGGTAAGACACCTCCCTCGACAC 2236
 QY 24 AepPheTrrThrLeuMetAepLysProLysValAlaIleLeuThrAlaGlyLeu 43
 Db 2235 CCGCGTCCGACCTGTATGCTCCCGCTCATGACATGAGTCTTACCGCGGTGGCAGC 2176
 QY 44 AlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGluIleAepPro 63
 Db 2175 TGCCCGGCTCAACGACCTCATTCGCGGCTCTGTAAAGCGAGCAGCAC---GAATTC 2119
 QY 64 SerIleGluIleIleCysTyrArgGlyTyrLysGlyLeuLeu-----LeuGlyAepSer 82
 Db 2118 GGCTACGAGTTCGTGGGATCGAGAACGCTACATGGGGCTGTGGAGCCAGGCGTCCG 2059
 QY 83 TyrProValThrAlaGluValArgLysValAlaGlyValLeuGluArgPheGlyLys 102
 Db 2058 CACCCGCTCACCGAGAGACACCCG---GGCATCTCTCCACAG---GGCGGACAC 2008
 QY 103 ValIleGlyAepSerArgValLysLeuThrAepValLysAepCysValLysArgGlyLeu 122
 Db 2007 ATCTGGGACGCTC-----AACAGGCCAACCCCTTCAGTACGCGAC 1963
 QY 123 ValLysGluGlyGluAepProGlnLysValAlaAepGlnLeuVal----- 138
 Db 1962 CGTAGAGATGGCAGCTGGGTGAGGAGCGAGCGTGCAGATCAGTCTGTCCGTCGAG 1903
 QY 139 LysAepGlyValAepIleLeuHicThrIleGlyLysAepPheThrAepThrAlaAla 158
 Db 1902 GAGCTGGGCTGAGACGGCTCATCCCGCGGCGGCGAGCGGCTGCATCCGCCAC 1843
 QY 159 AepLeuAlaAlaPheLeuAlaArgAsnAepTrrGlyLeuThrValIleGlyLeuProLys 178
 Db 1842 CGGCTCGT-----GAGAGGGGCTCAAGCTGTGGCTGTCCGAAG 1801
 QY 179 ThrValAepAepAepValPheProIleLysGlnSerLeuGlyAlaTrrThrAlaGlu 198
 Db 1800 ACCATCGACACGACCTGCGGAGACATCAGACTTCGCTTGACACCGCGCGCTC 1741
 QY 199 GlnGlyAlaArgTyrPheMetAepValAlaGluAsnAlaAepProAArgMcLLeu 218
 Db 1740 ATCGTCACGAGCGCTCGACCGCTGCACCTCACCGCGGAGCCACGACCGGGTGATG 1681
 QY 219 IleValHicGluValMetGlyArgAsnGlyTrrPleuThrAlaAlaThrAlaGlnGlu 238
 Db 1680 GTCGTG---GAGATCATGGCGCGCACCGCGCTTCCTCACCTG----- 1639
 QY 239 TyrArgLysLeuLeuAepArgAlaGluTrrPleuProGluLeuGlyLeuThrArgLys 258
 Db 1638 -----GAGAGCGCATCCCGGGGGCGG 1615
 QY 259 TyrGluValHicAlaValAepValProGluMetAlaIleAepLeuGluAlaGluAlaLys 278
 Db 1614 -----GACGTATCCTGATTCGAGAGATTCGCTACAGGTGAGTCCGTGTGAG 1564
 QY 279 ArgLeuArgGluValMetAepLysValAepCysValAepIlePhe---ValSerGluGly 297
 Db 1563 AAGATCCGCGCGCTCCACCGCGCTCGACCTTCATATGCGCATTCCTCGAAGGC 1504
 QY 298 Ala-----GlyValGluAlaIle----- 303
 Db 1503 GCGTTCGCCAGGCGGCGGAGCTGCGTGTGACACGCGGAGGCAATTCGCGCGCG 1444
 QY 304 ---ValAlaGluMetGlnAlaLysGlyGln----- 312

DB 1443 GCGGTGTCGGCTGCGCGGCTTCGGGAAAGCGCGTGGCCGACCTGCTGCGCCGACATC 1384
QY 313 -----GluValProArgAspAlaPheGlyHisIleLeu----- 323
DB 1383 GAGGGGAGATTCGGGTGACGATGCTGGGCACTCCACGCGGGGCGACCCCGACGCGC 1324
QY 324 -----LeuAspAlaValAsnProGly 330
DB 1323 GCGGACCGGTACTGCGCCTGCTACGCGCTCAAGGTGTGAGCTGTGAGCGCGGCG 1264
QY 331 LysTrp 332
DB 1263 CAGTGG 1258

RESULT 13
US-08-956-171E-312
Sequence 312, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunach
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44M storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 312:
SEQUENCE CHARACTERISTICS:
LENGTH: 7972 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 312:
US-08-956-171E-312

Alignment Scores:
Score: 6.18e-16 Length: 7972
Percent Similarity: 228.00 Matches: 102
Best Local Similarity: 42.248 Conservative: 75
Best Local Similarity: 24.348 Mismatches: 147
Query Match: 9.988 Indels: 95
DB: 4 Gaps: 19

US-10-701-200-6 (1-437) x US-08-956-171E-312 (1-7972)
QY 26 TyrThrLeuMetAsnLysProLysLeuValAlaIleLeuThrAlaGlyLeuAlaPro 45
DB 3647 TATGTCGTCATG-----AAGAAATTCAGATTAACTAGTGGAGATTCACCT 3697
QY 46 CysLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGluIleAspProSerIle 65
DB 3698 GGAATGAATGCTGCCGTAAAGACAGATTGTCGT-----ACGCAATTATTAATGAATTT 3751
QY 66 GluIleIleCysTyrArgGlyGlyTyrIleGlyLeuLeuGlyAspSerTyrProVal 85
DB 3752 GAGCTTATGCTGCTGATCATGCTTATCCAGAGATTGTTAAATGATGATTCATTAACCTT 3811
QY 86 ThrAlaGluValArgLysLeuAlaGlyValLeuGlnArgPheGlyGlySerValIleGly 105
DB 3812 -----GAATTAGATTCAGTTGGGATGATGATTCAGCT---GGAGTTACATTCCTGAT 3862
QY 106 AsnSerArgValLysLeuThrAsnValLysAspCysValLysArgGlyLeuValLysGlu 125
DB 3863 TCAGCAGAG-----TGTCCAGAG-----TTTAAGAG 3889
QY 126 GlyIleAspProGlnLysValAlaAlaAspGlnLeuValLysAspGlyValAspIleLeu 145
DB 3890 ---CAAGAAGTACGTAAGCTTGAATCGAATAAATCTTAAAGAGGATTCAGGCTT 3946
QY 146 HisThrIleGlyGlyAspAspThrAsnThrAlaAlaAspLeuAlaAlaPheLeuAla 165
DB 3947 GTAGTTATTTGGTGGTACGATGATTCGCGGTGCAACAGCATGAGAGATGTAA 4006
QY 166 ArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLysThrValAspAsnAspValPhe 185
DB 4007 GAA-----ATTCAAACTATCGTATTCCTGTGATGACATTCAGATGATATTCAT 4054
QY 186 ProIleLysGlnSerLeuGlyValAspThrAlaAlaGlnGlnGlyAlaArgTyrPheMet 205
DB 4055 GGTACTGATTTTACATTCATTCGATTCGACACAGCATTAATACGATTAATGCTTAATCAG 4114
QY 206 AsnValValAlaGluAsnAsnAlaAsnProArgMetLeuIleValHisGluValMetGly 225
DB 4115 AAATTTAGATGATTCGCTCAAGTCACGACGACGACATTTATTCAT---GAAGCAATGGGC 4171
QY 226 ArgAsnGlyGlyTyrPheLeuThrAlaAlaThrAlaGlnGlnTyrArgLysLeuLeuAspArg 245
DB 4172 CGTGAATGTGGA-----GATCTTA 4189
QY 246 AlaGluTyrPheLysProGlnLeuGlyLeuThrArgLysSerTyrGluValHisAlaValPhe 265
DB 4190 GCATTATGG-----GCTGGATTA-----TCAGTTGGTCCGACGACATTCGA 4231
QY 266 ValProGlnMetAlaIleAspLeuGlnAlaGluAlaLysArgLeuArgGluValMetAsp 285
DB 4232 GTTCAGAAAGTGAAGAAACAGAT-----ATTAAGAAATTAAGTGTAT 4270
QY 286 LysValAspCysValAsnIlePheValSerGluGlyAlaGlyValGlu----- 301
DB 4271 AAAATTGAA-----CAAGGATTTAAACGTGTGAAGAA 4303
QY 302 -----AlaIleValAlaGlnMetGlnAlaLysGlnGlnGluValProArgAspAla 318
DB 4304 CACTCAATCGTTCTTGTGACGAAAGGTGTATGACTGCGCAAGATTCGCAAAAAGATTA 4363
QY 319 PheGlyHisIleLysLeuAspAlaValAsnProGlyLysTrpPheGly-----GluGln 336
DB 4364 TCACAATACATCAATGTTGAT-----AATAGAGTGTCTGTGTTAGTCACGCTCAACGT 4417
QY 337 PheAlaGlnMetIleGlyAlaGlnLysThrLeuValGlnLysSerGlyTyrPheAlaArg 356
DB 4418 GGTGTGACCCCAACAGCGGTGATGAGATTTTATGATCATCGTTTATGTTGATGATGCG--- 4474
QY 357 AlaSerAlaSerAsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaVal 376
DB 4475 -----GTACACTTATTAATG 4489

QY 377 GluCyAlaPheArgGluSerGlyValIleGlyHisAspGluSerPheVal 396
 DB 4490 CAAGGTGAACACCTAAGGCTGTGAAT-----AAGAACATTA 4531
 QY 397 LeuArgAlaIleGluPheProArgIleGlyGlyPheProPheAsnIleAspThrAsp 416
 DB 4532 ATTGTAGCAACATCTTTTGTATGAATTTTGTATGAATCATTAATTTGATTATAGT 4591
 QY 417 -TTPheAsnSerMetLeuSerGluIleGlyInProGlyGlyValGlu 434
 DB 4592 CTATATGAACTTGCTAACAAGATTATATTAAGATTTCAGAGGAATTATTA 4646
 RESULT 14
 US-08-781-986A-312
 Sequence 312, Application US/08781986A
 Patent No. 6737248
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5255
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781, 986A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Benson, Bob
 REGISTRATION NUMBER: 30,446
 REFERENCE/DOCKET NUMBER: PB248PP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 312:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7972 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-781-986A-312
 Alignment Scores:
 Pred. No.: 6.18e-16 Length: 7972
 Score: 228.00 Matches: 102
 Percent Similarity: 42.24% Conservative: 75
 Best Local Similarity: 24.34% Mismatches: 147
 Query Match: 9.98% Indels: 95
 DB: 4 Gaps: 19
 US-10-701-200-6 (1-437) x US-08-781-986A-312 (1-7972)
 QY 26 TyrThrLeuMetCysProGlyValIleGlyHisAspGluSerPheVal 45
 DB 3647 TATGTCCTCATG-----AAGAAATTCAGATTTAAGTGTGAGATTTCACCT 3697
 QY 46 CysLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGluIleAspProSerIle 65
 DB 3698 GGAATGATCTGCTGCGTAAGACGATTGCTG-----ACAGCAATTACAAATGAAT 3751

QY 66 GluIleIleCysTyrArgGlyIleGlyValIleGlyLeuLeuGluValAspSerTyrProVal 85
 DB 3752 GAAGTTATGCTGTATCATCTGTTACCAAGATTGTTAATGATGATATTAACCTT 3811
 QY 86 ThrAlaGluValArgIleGlyValIleGlyValIleGlyValIleGlyValIleGly 105
 DB 3812 -----GAATTAGATCAAGTTGGGATACGATTCAGCT--GAGGTACATCTGTAT 3862
 QY 106 AsnSerArgValIleValThrAsnValIleAspCysValIleValIleValIleVal 125
 DB 3863 TCAGCAAGA-----TGCCAGAG-----TTTAAGAG 3889
 QY 126 GlyGluAspProGluValIleValIleValIleValIleValIleValIleVal 145
 DB 3890 ---CAAGATGACATTAAGTTGCAATCGAAACCTTACGTAAAGAGGATTGAGGCTT 3946
 QY 146 HisThrIleGlyIleValAspThrAsnThrAlaIleValIleValIleValIleVal 165
 DB 3947 GTAGTTATGCTGTATCATCTGTTATCCGCTGCAACGCAATGAGAAATGTA 4006
 QY 166 ArgAsnAsnTyrGlyLeuThrValIleGlyLeuProGlyThrValIleValIleVal 185
 DB 4007 GAA-----ATCAACTATCGGATTCCTGTCAGATTCATGATATCAAT 4054
 QY 186 ProIleValSerLeuGlyValIleValIleValIleValIleValIleValIleVal 205
 DB 4055 GGTACTGATTTTACATTTGATTTGACACAGCATTAATTCATTTATGCTTATGCTAC 4114
 QY 206 AsnValValIleGluAsnAsnAlaAsnProGlyMetLeuIleValIleGluValIleVal 225
 DB 4115 AAATTAAGATATCTGCTCACTGCAAGCAAGCAATTTATCAAT--GAAGCAATGCGC 4171
 QY 226 ArgAsnCysGlyTyrLeuThrAlaIleValIleGluValIleValIleValIleVal 245
 DB 4172 CGTATTTGCA-----GATCTA 4189
 QY 246 AlaGluTyrLeuProGluLeuGlyLeuThrArgGluSerTyrGluValIleValIleVal 265
 DB 4190 GCATTTATG-----GCTGATTA-----TCAGTGTGCTGAGCAATGCTA 4231
 QY 266 ValProGluMetAlaIleAspGluValIleGluValIleValIleValIleValIleVal 285
 DB 4232 GTTCCAGAGTGAACAGAT-----ATTAAAGAAATGCTGAT 4270
 QY 286 IysValAspCysValAsnIlePheValSerGluValIleGlyValIleValIleVal 301
 DB 4271 AAAATTGA-----CAAGTATTAACGCTGTAAGAA 4303
 QY 302 -----AlaIleValIleGluMetGlnAlaIleGlyGlnValIleProArgAspAla 318
 DB 4304 CACTCAATGCTTCTGTATGAGAGAGGTTGTATGATCTGCGCAAGATTGTCAAAAAGATTA 4363
 QY 319 PheGlyHisIleIleValAspAlaValAsnProGlyValIleThrPheGly-----GluGln 336
 DB 4364 TCACAATACATCAATGTGAT-----AATAGTGTCTGTATGATGTCACGTTCAACGT 4417
 QY 337 PheAlaGluMetIleGlyValIleGluValIleValIleValIleValIleValIleVal 356
 DB 4418 GGTGTGATGCCCAACAGGTGCGTATGATTTTGACATCAGCTTATGATGATATGCG-- 4474
 QY 357 AlaSerAlaSerAsnValAspAspMetArgLeuIleIleSerCysAlaIleValIleVal 376
 DB 4475 -----GTAGCTTATTAATG 4489
 QY 377 GluCyAlaPheArgGluSerGlyValIleGlyHisAspGluSerPheVal 396
 DB 4490 CAAGGTGAACACCTAAGGCTGTGAAT-----AAGAACATTA 4531
 QY 397 LeuArgAlaIleGluPheProArgIleGlyGlyPheProPheAsnIleAspThrAsp 416
 DB 4532 ATTGTAGCAACATCTTTTGTATGAATTTTGTATGAATCATTAATTTGATTATAGT 4591
 QY 417 -TTPheAsnSerMetLeuSerGluIleGlyInProGlyGlyValGlu 434

```

DB      4592 CTAATGAACTGCTACAGATTATCTATATTAAGATTTCAGAGGAAATTATATAA 4646
RESULT 15
US-09-489-039A-5087
/ Sequence 5087, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2004001
/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 5087
/ LENGTH: 1017
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5087

```

```

Alignment Scores:
Pred. No.:      3,166-17      Length:      1017
Score:          227.00      Matches:      105
Percent Similarity: 38.84%      Conservative: 62
Best Local Similarity: 24.42%      Mismatches: 136
Query Match:      9.94%      Indels:      128
DB:              4          Gaps:      18

```

US-10-701-200-6 (1-437) x US-09-489-039A-5087 (1-1017)

```

QY      13  ILeaTgPhaCyehiEtrPhePheLeuAenPheAenPheTyThrLeuWeLAsnLySPro 32
DB      22  ATTATATTTTGC-----TTCTAAGTTCAAGAGTAGCATGAT-----60
QY      33  LySLeValAlAlLeuThrAlaGlyLeuAlaProCyLeuAenSerAlaIleGly 52
DB      61  AAGAAATGCGGTGTGTGACAAAGTGGCGGTGATGCGCGGCGGTGAACGACGAAATTGCT 120
QY      53  SerLeuIleGluArgTyThrGluLeuAProSerIleGluIleIleCySyrArgGly 72
DB      121 GGGCGTTGTCGC-----GGCGATTACGGAAAGCTCGAACTTTGGAACTCTATGAC 174
QY      73  GlyTyLeuGlyLeu-----LeuLeuGlyAAspSerTyProValThrAla 87
DB      175 GGTACCTCGGATTTGATGAAAGACCGTATGCTTACACTGACCGCTTACAGCGTTCCGAC 234
QY      88  GluValArgLysAlaGlyValLeuGlnArgPheGlyGlySerValIleGlyAenSer 107
DB      235 ATGATTAACCGC-----GGCGGTAACCTTCTCGGCGTCCGCT 270
QY      108 ArgValLysLeuThrAenValLysAAspCyValLysArgGlyLeuValLysGluGly 127
DB      271 CGC-----TTCCCGGAATTCGCGAA-----GAA 294
QY      128 AAspProGlnLysValAlaAlaAAspGlnLeuValLysAAspGlyValAAspIleLeuHisThr 147
DB      295 CACATCGCGCGTGTGCTATCGAAACATGAAAGACGCGGCTGTGACCGCGTGTG 354
QY      148  ILeGlyGlyAAspAAspThrAenThrAlaAlaAlaAAspLeuAlaAlaPheLeuAlaArgAen 167
DB      355 ATCGGCGGTGACGCTTCTATATGAGGCGATGCGCGCTG-----ACC 396
QY      168 AAsnTyGlyLeuThrValIleGlyLeuProLysThrValAAspAAspValPheProIle 187
DB      397 GAGATGGGCTTCCCATGCAATCGGCTGCGGCGGACCATGACACGATATCAAGGCACT 456
QY      188  LySLeuSerLeuGlyValAAspThrAlaAlaGluGlnGlyAlaArgTyRPhenLeuAenVal 207
DB      457 GACTACACATGCGGCTTCTTACTGCGCTGAGACCGGTGTGGAACGATTCACGCTTTC 516

```

```

QY      208  ValAlaGluAenAenAlaAAsnProArgMetLeuIleValHisGluValMetGlyArgAen 227
DB      517 CGCGACACCTCTCTTCGCACACGATATCTCCGTGGT--GAAGTAGTGGCGGTTAC 573
QY      228  CySGlyTyRPhenThrAlaAlaThrAlaGlnGlyTyArgLysLeuLeuAAspArgAlaGlu 247
DB      574 TGTGGCAACCTGACCTCGGCGGCGGCG-----ATTGCGGCGGCGGTGTAG 618
QY      248  TrpLeu-----ProGluLeuGlyLeuThrArgGlySerTyRGlValHisAlaValPhe 265
DB      619 TTCATCATGTCGTCGAAAGTGAAATATACCGCTGAC-----654
QY      266  ValProGluMetAlaAlaAAspLeuGluAlaGlyValLys-----278
DB      655 -----GATCTGGTGCAGAAATCAAAAGCCGGTATCGCGAAAGGAAA 696
QY      279 -----ArgLeuArgGluValMetAAspLysValAAspCyValAAsnIlePhe 293
DB      697 AAACACGCTATCGTGGCCATCCGACGACATGTGCGACGTTGACGAGCTGGCAAGCTAC 756
QY      294  ValSerGluGlyAlaGlyValGlu-----AlaIleValAlaGluMetGlnAlaLys 310
DB      757 ATCGAGAAAGAGACTGCGCGTGAACCTCGCGCACCGTCTCGGCACATTCACGCGCGC 816
QY      311  GlyGlnGluValProArgAAspAlaPheGlyHisIleLysLeuAAspAlaValAAsnProGly 330
DB      817 GGTTCGCCGCTTCTTACGAC-----837
QY      331  LysTrpPheGlyGluGlnPheAlaGlnMetIleGlyAlaGluLysThrLeuValGlnLys 350
DB      837 -----837
QY      351  SerGlyTyRPhenAlaArgAlaSerAlaSerAenValAAspAAspMetAAspLeuIleLysSe 370
DB      838 -----CGCATCTGCGCTTCCCGCATGCGGCGGCGCTTATGCGATGAGCTGTG 882
QY      370  rCyAlaAAspLeuAlaValGluCyAlaPheArgArgGlySerGlyValIleGlyHisAs 390
DB      883 CTGCGAGGCGCATGGGCGGCG- CGTTGCGTGGCATCCAGAAAGACGAGTGTGACACGCA 941
QY      390  pGluAAspAAsnGluAenValLysArgAlaIleGluPheProArgIleLysGlyLysPr 410
DB      942 C-----ATCATGATGCCATGAGAAACATGAAG-----CGTCC 974
QY      410  pPheAAsnIleAAspThrAAspTrpPheAAsn 419
DB      975 GTTCAAG-----AACGACTGCTGTGAT 996

```

Search completed: March 1, 2005, 17:14:53
 Job time : 265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2005, 17:01:32 ; Search time 749 Seconds
(without alignments)
3457.177 Million cell updates/sec

Title: US-10-701-200-6
Perfect score: 2284
Sequence: 1 DVVTMPHYLTADIRCFWFF.....FNSMSEIGQPKGKVEVSH 437

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-DB=PubMed Applications NA -QFMT=faaCap SURF1=trmbp -MIMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEATSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10701200 @CGN 1.1480 @runat 25022005 102834_10410
-NUPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCCK=100
-LONGLOG -DEV TIMEOUT=120 -MAXN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FAPOP=6 -FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : 3 Published Applications NA:*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2284	100.0	1311	9 US-09-934-901-15	Sequence 15, Appl
2	2284	100.0	1311	9 US-09-934-868-5	Sequence 5, Appl
3	2284	100.0	1311	10 US-09-941-947A-1	Sequence 1, Appl
4	2284	100.0	1311	10 US-10-320-924-15	Sequence 15, Appl
5	2284	100.0	1311	15 US-10-320-874-15	Sequence 1, Appl
6	2284	100.0	1311	17 US-10-363-567-1	Sequence 15, Appl
7	2284	100.0	1311	18 US-10-321-210-15	Sequence 15, Appl
8	659	28.9	552	17 US-10-369-493-38035	Sequence 38035, A
9	306.5	13.4	1050	17 US-10-369-493-31795	Sequence 31795, A
10	298	13.0	9025608	15 US-10-156-761-1	Sequence 1, Appl
11	294	12.9	1026	15 US-10-156-761-6660	Sequence 6060, Ap
12	293	12.8	1570	17 US-10-425-114-9204	Sequence 9204, Ap
13	293	12.8	1765	17 US-10-425-114-6293	Sequence 6293, Ap
14	293	12.8	1945	17 US-10-425-114-10982	Sequence 10982, A
15	293	12.8	4413	17 US-10-424-599-141147	Sequence 141147,
16	284.5	12.5	1023	15 US-10-156-761-7102	Sequence 7102, Ap
17	284.5	12.5	9025608	15 US-10-156-761-1	Sequence 1, Appl
18	279.5	12.2	1998	18 US-10-437-963-101981	Sequence 101981,
19	270	11.8	1347	17 US-10-282-122A-11067	Sequence 11067, A
20	269.5	11.8	1023	15 US-10-156-761-2809	Sequence 2809, Ap
21	269.5	11.8	2100	18 US-10-767-701-14854	Sequence 14854, A
22	267.5	11.7	3405	18 US-10-425-115-165105	Sequence 165105,
23	263.5	11.5	2310	18 US-10-437-963-12274	Sequence 12274, A
24	262	11.5	960	17 US-10-369-493-26592	Sequence 26592, A
25	261	11.4	1035	17 US-10-369-493-42948	Sequence 42948, A
26	259	11.3	1358	18 US-10-437-963-6352	Sequence 6352, Ap
27	258.5	11.2	2663	17 US-10-425-115-121483	Sequence 121483,
28	255	11.1	2149	17 US-10-424-599-128266	Sequence 128266,
29	254	11.1	956	17 US-10-369-493-33339	Sequence 33339, A
30	249.5	10.9	2123	18 US-10-425-115-119704	Sequence 119704, A
31	248	10.9	1773	17 US-10-369-493-33356	Sequence 33356, A
32	246.5	10.8	2664	17 US-10-369-493-45685	Sequence 45685, A
33	246.5	10.8	2664	18 US-10-793-639-404	Sequence 404, App
34	242.5	10.6	2024	17 US-10-424-599-68225	Sequence 68225, A
35	242.5	10.6	2024	17 US-10-425-114-11330	Sequence 11330, A
36	242.5	10.6	1441	17 US-10-425-114-10007	Sequence 10007, A
37	241.5	10.6	2812	17 US-10-424-599-115967	Sequence 115967, A
38	241.5	10.6	960	17 US-10-282-122A-36804	Sequence 36804, Ap
39	240.5	10.5	963	9 US-09-815-242-9687	Sequence 9687, Ap
40	240.5	10.5	963	17 US-10-282-122A-38849	Sequence 38849, A
41	240.5	10.5	963	17 US-10-282-122A-39841	Sequence 39841, A
42	240.5	10.5	3661	18 US-10-437-963-1690	Sequence 1690, Ap
43	239.5	10.5	963	9 US-09-815-242-6325	Sequence 6325, Ap
44	238.5	10.4	963	17 US-10-369-493-47325	Sequence 47325, A
45	238.5	10.4	963	17 US-10-369-493-47325	Sequence 47325, A

ALIGNMENTS

RESULT 1
US-09-934-901-15
; Sequence 15, Application US/09934901
; Patent No. US20020110885A1
; GENERAL INFORMATION:
; APPLICANT: Kofifae, Matheos
; APPLICANT: Odom, J. Martin
; APPLICANT: No. US20020110885A1ton, Kelley C.
; TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: C1619 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229, 906
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office. 97
; SEQ ID NO 15
; LENGTH: 1311
; TYPE: DNA

! ORGANISM: METHYLOMONAS SP.
US-09-934-901-15

Alignment Scores:

Pred. No.:	4,19e-253	Length:	1311
Score:	2284.00	Matches:	437
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-701-200-6 (1-437) x US-09-934-901-15 (1-1311)

```

QY 1 AapValValThThrProTyRHisLeuThraAlaAapIleArgPheCyHisTrpPhePhe 20
DB 1 GATGGGTGACATGGCCCTATCACTTAAGCGCTGATTAATTCATTTGTCATGGTTTTTT 60
QY 21 LeuAenPheAenPheTyRThrLeuMetAenLySProLySLeuValAlaIleLeuThra 40
DB 61 CTTAACTTTAACTTTCAACGCTCATGAACAACTTAATAAAAGTTGCAATCTGACAGCA 120
QY 41 G1yG1yLeuAlaProCySLeuAenSerAlaIleG1ySerLeuIleG1uArgTyRThrg1u 60
DB 121 GGCGGCTGGCCCTGTTGTAATTCGCAATCGGTAAGTTTGAATCGAACGTAATACGNA 180
QY 61 IleAapProSerIleG1uIleIleCyS1yRArgG1yTyR1yG1yLeuLeuLeuG1y 80
DB 181 ATCGATCTTACGATGAATAATCATTTGCTATCGCGGGGTTATTAAGGCTGTTGCTGGGC 240
QY 81 AapSerTyRProValThraIleValArgLySLeuAlaG1yValLeuG1uArgPheG1y 100
DB 241 GATTCCTTATCCAGTAAAGCCGCAAGCGCTTAATAAAAGCGGGCTTCTGCAACGTTTGGC 300
QY 101 G1ySerValIleG1yAenSerArgValLySLeuThraAenValLyAapCyValLySArg 120
DB 301 GGTTCTGTATCGGCAACGCGCGCTCAATATGACCAATGTCAAAAGACTGCGTAAGC 360
QY 121 G1yLeuValLyG1uG1yG1uAenProG1uLySValAlaAlaAapG1uLeuValLySap 140
DB 361 GGTGGTGGTCAAGAGGGGTAAAGATCGCAAAAGTCGCGGCTGATCAATTTGTTAAGAT 420
QY 141 G1yValAapIleLeuHisThrIleG1yG1yAapAapThraAenThraAlaAlaAapLeu 160
DB 421 GGTCGATATCTGCAACCAATCGCGCGCATGATACCAATACGCGACAGCGATTTTG 480
QY 161 AlaAlaPheLeuAlaArgAenAenTyRg1yLeuThraValIleG1yLeuProLySThra 180
DB 481 GCAGCATTCCTGGCCGAATAATTAATTAACGATGACGCTCATTTGTTAAACCGTC 540
QY 181 AapAenAapValPheProIleLySLeuG1uArgTyRThraAlaIleG1uG1uG1y 200
DB 541 GATTAACGAGTATTTCCGATCAAGCATCACTAGGTGCTTGACCTGCCCGCAAGCAGC 600
QY 201 AlaArgTyRPhemeAenValAlaIleG1uAenAenA1AapProArgMetLeuIleVal 220
DB 601 GCGCTTATTTATGATCAAGCTGGTGGCCGAATAACCAAGCCCAACGATGCTGATCGTA 660
QY 221 HisG1uValMetG1yArgAenCySg1yTyRLeuThraAlaIleThraIleG1uLySArg 240
DB 661 CACGAAGTATGGGCGGTAACTGGCGCTGACGCGTGAACCGCGCAAGATATGCGC 720
QY 241 LySLeuLeuAapArgAlaIleUtrPleuProG1uLeuG1yLeuThraArgLySArgTyR 260
DB 721 AATTAATCTGGACCGTGCCAGGTGGTGGCGGAATGGGTTGACTCGTAATCTTAATGA 780
QY 261 ValHisAlaValPheValProG1uMetAlaIleAapLeuG1uAlaG1uAlaLySArgLeu 280
DB 781 GTGCAACGCGGTATTCCTCCGAATGGGATCGATCGACTGAAGCCGAAGCCAAAGCCGCT 840
QY 281 ArgG1uValMetAapLySValAapCySValAenIlePheValSerG1uG1yAlaG1yVal 300
DB 841 CCGGAAGTATGAGCAAAAGTCATTTGCGTCAACATCTTCTGTTCCGAAGTCCCGGCGTC 900

```

```

QY 301 G1uAlaIleValAlaG1uMetG1uAlaLySLeuG1uG1uValProArgAapAlaPheG1y 320
DB 901 GAAGCTATCGTCGCGGAATGAGGCGCAAGCGCAAGATGCCCCGCGATGCTTGGC 960
QY 321 HisIleLySLeuAapAlaValAapProG1yTyRTrpPheG1yG1uG1uPheAlaIleMet 340
DB 961 CACATCAAACTGAGATGGCGTCAACCTGTAATGTTCCGGAGCAATTCGCGCAAGTTG 1020
QY 341 IleG1yAlaG1uLySThraLeuValG1uLySArgTyRThraAlaArgAlaSerAlaSer 360
DB 1021 ATAGCGCGGAAAAAACCCCTGTACAAAAATCGGGATCTTCGCGCGCTTCTCTCC 1080
QY 361 AenValAapAapMetArgLeuIleLySArgTyRValAapPheValAlaG1uCySAlaPhe 380
DB 1081 AACGTTGACGACATCGCTTATCAATCGTGCGCGCATGCGCGTGAAGTCCGCTTC 1140
QY 381 ArgArgLySArgTyRValIleG1yHisAapG1uAapAenG1yAenValLeuArgAlaIle 400
DB 1141 CCGCGCGAGTCTGGGTATGCGTCAACGACGAACCAACGCGAAGCTGTGCTCGATC 1200
QY 401 GluPheProArgIleLyG1yLySProPheAenIleAapThraAapTrpPheAenSer 420
DB 1201 GAGTTCCGCGCATCAAGGCGCGCAACCGTTCAATATCGACACCGACTGTTCAATAGC 1260
QY 421 MetLeuSerG1uIleG1yG1uProLySLeuG1yLySValG1uValSerHis 437
DB 1261 ATGTTGACGGAATCGCCACGCTTAAGCGGTAAAGTCGAAGTCAGCCAC 1311

RESULT 2
US-09-934-868-5
; Sequence 5, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Kofias, Matheos
; APPLICANT: Odom, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: C14596 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229, 858
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: METHYLOMONAS SP.
US-09-934-868-5

Alignment Scores:
Pred. No.: 4,19e-253 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-701-200-6 (1-437) x US-09-934-868-5 (1-1311)
QY 1 AapValValThThrProTyRHisLeuThraAlaAapIleArgPheCyHisTrpPhePhe 20
DB 1 GATGGGTGACATGGCCCTATCACTTAAGCGCTGATTAATTCATTTGTCATGGTTTTTT 60
QY 21 LeuAenPheAenPheTyRThrLeuMetAenLySProLySLeuValAlaIleLeuThra 40
DB 61 CTTAACTTTAACTTTCAACGCTCATGAACAACTTAATAAAAGTTGCAATCTGACAGCA 120
QY 41 G1yG1yLeuAlaProCySLeuAenSerAlaIleG1ySerLeuIleG1uArgTyRThrg1u 60
DB 121 GGCGGCTGGCCCTGTTGTAATTCGCAATCGGTAAGTTTGAATCGAACGTAATACGNA 180
QY 61 IleAapProSerIleG1uIleIleCyS1yRArgG1yTyR1yG1yLeuLeuLeuG1y 80

```

```

Db      181 ATCGATCTAGCAATAATCTTCTATCGCGCGGTTAAAGCCGTGTGCGGCG
Qy      81 AapSerTyProValThrIaGluValArgLyLeuAlaGlyLeuGlnArgPheGly 100
Db      241 GATTCTTATTCCTAATACGGCCGGAAGTGCTGTAAGAGCGGGTCTTCTGCAACCTTTTGGC 300
Qy      101 GlySerValIleGlyAnSerArgValIleLeuThrAsnValIleAspCyValIleValArg 120
Db      301 GGTTCGTGATCGGCAACAGCCGCGTCAATTGACCAATGTCGAAGACTGCGTGAACCGC 360
Qy      121 GlyLeuValIleGlyGluGlyLeuAspProGlnIleValAlaIleAspGlnLeuValIleAsp 140
Db      361 GGTTCGTGCAAGAGGCTGAAGATCCGCAAAAAGTCGCGCTGATCAATGGTTAAGAT 420
Qy      141 GlyValAspIleLeuIleThrIleGlyIleAspAspThrAsnThrAlaIleAspLeu 160
Db      421 GGTTCGTGATATTCGACACCATCGCGCGGATGATACCAATACGCGACGCGGATTTG 480
Qy      161 AlaIlePheLeuAlaArgAsnAsnTyGlyLeuThrValIleGlyLeuProLyThrVal 180
Db      481 GCGACATTCCTGCGCAAGAATATTAACGACTGACCGTACTTGTCTTACTTAACCGTC 540
Qy      181 AspAsnAspValPheProIleTyGlyInSerLeuGlyAlaIleThrAlaIleGluGlnGly 200
Db      541 GATTAACGACGATATTCGATCAAGCAATCACTAGGCTTGATGACTCGCGCGCAAGGC 600
Qy      201 AlaArgTyPheMetAsnValIleAlaGluAsnAsnAlaAsnProLyMetLeuIleVal 220
Db      601 GCGCGTATATTCATGAAACGTGGTGCGCGAAGAACGCAACCCACGACATGCTGATCGTA 660
Qy      221 HisGluValMetGlyArgAsnCyGlyTyPLeuThrAlaIleThrAlaGlnGluTyArg 240
Db      661 CACGAAGTATGAGCGCGCTAATCGCGCTGCTACCGCTGCAACCGCGCAAGATATCCG 720
Qy      241 LyLeuLeuAspArgAlaGluTyPLeuProGlnLeuGlyLeuThrArgGlySerTyGly 260
Db      721 AATTAATGACCGCTGCGCGAATGCTGCGGAAATGGGTTGACTGCTGATCTATGAA 780
Qy      261 ValHisAlaValPheValProGlnMetAlaIleAspLeuGlnAlaGluIleValArgLeu 280
Db      781 GTTCACGCGGATATTCGTCCGGAATAGCGCATGACCTGGAACCGCAAGCAAGCCCTCG 840
Qy      281 ArgGluValMetAspLyValIleAspCyValAsnIlePheValSerGluGlyAlaGlyVal 300
Db      841 CCGGAAGTATGACCAAGCTCGATTCGCTGCAATCTTCTGTTCCGAAGGTGCGGCGTC 960
Qy      301 GluAlaIleValAlaGlnMetGlnAlaIleGlyGlnGluValProArgAspAlaPheGly 320
Db      901 GAAGCATATCGTCGCGAATGACAGGCCAAGCCAGAAAGTGCAGCGCATGCTCGGC 960
Qy      321 HisIleIleLeuAspAlaValAsnProGlyLyTyPLeuGlyGluGlnPheAlaGlnMet 340
Db      961 CACATCAAACTGATCGGTCAAACCTGCTGAATGCTGCGGACGAGCAATTCGGCGCAGATG 1020
Qy      341 IleGlyAlaGluTyPLeuThrLeuValIleGlnIleSerGlyTyPLeuAlaArgAlaSer 360
Db      1021 ATTAGCGCGCGAAGAAACCTCGTGAACAAATCCGGATACCTTCCCGCTGCTTCTGCTTC 1080
Qy      361 AsnValAspAspPheArgLeuIleIleSerCyAlaAspLeuAlaValGluCyAlaPhe 380
Db      1081 AACGTGACGACATGCGTTGATCAATCGTGCAGCACTTGCGCGGTGCGAGTCCGCTTC 1140
Qy      381 ArgArgGluSerGlyValIleGlyIleAspGlyAspAsnGlyAsnValIleValAlaIle 400
Db      1141 CCGCGGAGCTCGCGTATCGGTCACGACGAAAGACACCGCAACGTTGCTGCGGATC 1200
Qy      401 GluPheProArgIleTyGlyGlyTyPLeuProPheAsnIleAspThrAspTyPLeuAsnSer 420
Db      1201 GATTTCCCGGCAATCAAGGCGCGCAACCTTCAATATGCAACCGACTGTTCAATAGC 1260
Qy      421 MetLeuSerGluIleGlyGlnProLyGlyIleValAlaGluValSerHis 437

```

```

Db      1261 ATGTTGAGCGAAATCGGCCACCTTAAGCGGTTAAAGTCGAAGTCAGCAC 1311
RESULT 3
US-09-941-947A-1
; Sequence 1, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
; APPLICANT: Brzoskiewicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: Dicostimo, Deana J.
; APPLICANT: Kofias, Matheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odom, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Rouverie, Pierre E.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: CL1903 US NA
; CURRENT APPLICATION NUMBER: US/09/941,947A
; CURRENT FILING DATE: 2001-09-01
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Methylobionas 16a
US-09-941-947A-1

Alignment Scores:
Pred. No.: 4,19e-253 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Query Match: 10 Gaps: 0

US-10-701-200-6 (1-437) x US-09-941-947A-1 (1-1311)
Qy      1 AapValAlaThrTrpProTyHisLeuThrAlaAspIleArgPheCyHisIleTrpPhe 20
Db      1 GATGTGTGACATGAGCCCTATCACTTAACGGCTGATATTCGATTTGTGATTTGTTT 60
Qy      21 LeuAsnPheAsnPheTyThrIleuMetAsnLyProLyIleValAlaIleLeuThrAla 40
Db      61 CTTAACCTTAACCTTCAACGCTCATGAACAAACCTAAAGAGTTGCAATCACTGACAGCA 120
Qy      41 GlyGlyLeuAlaProCyLeuAsnSerAlaIleGlySerLeuIleGluArgTyThrGly 60
Db      121 GCGCGCTTGGCGCTTGTGTAATCCGCATCGGATCGGATGTTGATCGAAGCTTATCCGAA 180
Qy      61 IleAspProSerIleGluIleIleCyTyTyArgGlyTyTyTyGlyLeuLeuLeuGly 80
Db      181 ATCGATCTTAGCATTAAGAAATCTTGTCTATCGCGCGGTTATTAAGCCCTGTTGCTGGC 240
Qy      81 AapSerTyProValThrIaGluValArgLyLeuValIleGlyValLeuGlnArgPheGly 100
Db      241 GATTCTTATTCAGTAAACCGCCGAAAGTCGTAAAGAGCGGGTGTTCGAAACGTTTGGC 300
Qy      101 GlySerValIleGlyAnSerArgValIleLeuThrAsnValIleAspCyValIleValArg 120
Db      301 GGTTCGTGATCGGCAACAGCCGCGTCAATTGACCAATGTCGAAGACTGCGTGAACCGC 360
Qy      121 GlyLeuValIleGlyGluGlyLeuAspProGlnIleValAlaIleAspGlnLeuValIleAsp 140
Db      361 GGTTCGTGCAAGAGGCTGAAGATCCGCAAAAAGTCGCGCTGATCAATGGTTAAGAT 420
Qy      141 GlyValAspIleLeuIleThrIleGlyIleAspAspThrAsnThrAlaIleAspLeu 160
Db      421 GGTTCGTGATATTCGACACCATCGCGCGGATGATACCAATACGCGACGCGGATTTG 480

```

QY	161	AlAlaIaPheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLyseThrVal	180
Db	481	GCAGCATTCCTCGGCAGAAATATTAACGACTGACCGCTCATTTGGTTTACCTAAACCGTC	540
QY	181	AspAsnAspValA1PheProIleIeysGlnSerLeuGlyA1aTrpThrA1aIaGluGlnGly	200
Db	541	GATAAACGACGATATTCGATCAACCAATCATTAAGTGTGTTGACTGCGCGCAGCAAGGC	600
QY	201	AlAArgTyrPheMetAsnValIValAlaGluAsnAsnAlaAsnProArgMetLeuIleVal	220
Db	601	GCGGTTATATTCATGAACAGTGGTGGCCGAAACACGCAACCCACGATCTCATGCTA	660
QY	221	HiGluValMetGlyArgAsnCGlyTTrpLeuThrAlaA1aTrpA1aGluGlnTyrArg	240
Db	661	CACGAAAGTGAATGGCCGTTAACTCGGCGCTGGCTTACCGCTGCAACCGCCACGAAATATCGC	720
QY	241	LyseLeuLeuAspArgAlaGluTrpLeuProGluLeuGlyLeuThrArgGluSerTyrGlu	260
Db	721	AAATTAATCGACCGCGCGAATGGTGGTCGGAATTTGGGTTTGACTCGTAATCTTAACAA	780
QY	261	ValHiSalAValPheValProGluMetAlaIleAspLeuGluAlaGluAlaLyseArgLeu	280
Db	781	GTGACACCGGATTCGTTCCCGGAAATGGCGATCACTGACCTGGAAGCCGCAACGCGCTGC	840
QY	281	ArgGluValMetAspLyseValAspCysAlaAsnIlePheValSerGluGlyAlaGlyVal	300
Db	841	CGCCAAAGTGAAGGCAAAAGTGAATTCGCTCAACATCTTCGTTCCGAAAGTGCGCGCGCTC	900
QY	301	GluValAlaIleValAlaGluMetGlnAlaLyseGlyGlnGluValProArgAspAlaPheGly	320
Db	901	GAACCTATTCGTCCGCGAATATCGAAGCCAAAGCCGCAAGAAAGTCCGCGCGATTCGCGC	960
QY	321	HiIleIeLyseLeuAspAlaValAsnProGlyLyseTrpPheGlyGluGlnPheAlaGlnMet	340
Db	961	CACATCAAACTGATGCGGTCAACCTCGGTAAATGTTCCGCGCGAGCAATTCGCCACAGATG	1020
QY	341	IleGlyAlaGluLyseThrLeuValGlnLyseSerGlyTyrPheAlaArgAlaSerAlaSer	360
Db	1021	ATAGACCGCGGAAAAAACCTCTGTACAAAATTCGGGATATCTTCGCCCGCTTCGTGCTTCC	1080
QY	361	AsnValAspAspMetCArgLeuIleLyseSerCysAlaAspLeuAlaValaGluCysAlaPhe	380
Db	1081	AACTTGAACGACATTCGTTTATCAAAATCGTGCGCCCACTTGGCGGTTCGATGGCGGCTTC	1140
QY	381	ArgArgGluSerGlyValaIleGlyAlaIleAspGluAspAsnGlyAsnValLeuArgAlaIle	400
Db	1141	CGCGCGAGAGTCTGGCGGTATCGGTCAACGACGAAAGCAACGCGCAACGTGTTCGTCGATC	1200
QY	401	GluPheProArgIleLyseGlyGlyLyseProPheAsnIleAspTrpThrAspTrpPheAsnSer	420
Db	1201	GAGTTTTCGCCCATCAACAGGCGGCAACCCGTTCAATATTCGACACCGATTCGATTAATGAC	1260
QY	421	MetLeuSerGluIleGlyGlnProLyseGlyGlyLyseValaGluValaSerHis	437
Db	1261	ATGTTGACGAAATTCGCCACGCTTAAGGCGGTAAATGTCGAAGTCAGGCCAC	1311
RESULT 4			
US-10-320-924-15			
Sequence 15, Application US/10320924			
Publication No. US20030129721A1			
GENERAL INFORMATION:			
APPLICANT: Kotfas, Matheos			
APPLICANT: Odom, J. Marcia			
APPLICANT: No. US20030129721A1ton, Kelley C.			
APPLICANT: Ye, Rick			
TITLE OR INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN			
FILE REFERENCE: CL1619 US NA			
CURRENT APPLICATION NUMBER: US/10/320,924			
CURRENT FILING DATE: 2002-12-17			
PRIOR APPLICATION NUMBER: US/09/934,901			
PRIOR FILING DATE: 2001-08-22			
PRIOR APPLICATION NUMBER: 60/229,906			
PRIOR FILING DATE: September 1, 2000			

: NUMBER OF SEQ ID NOS: 20			
: SOFTWARE: Microsoft Office 97			
: SEQ ID NO 15			
: LENGTH: 1311			
: TYPE: DNA			
: ORGANISM: METHYLOMONAS SP.			
US-10-320-924-15			
Alignment Scores:			
Pred. No.:			
Score: 4.19e-253			
Percent Similarity: 2284.00%			
Best Local Similarity: 100.00%			
Query Match: 100.00%			
DB: 15 Gaps: 0			
US-10-701-200-6 (1-437) x US-10-320-924-15 (1-1311)			
Qy	1	AspValValThrTrpProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhe	20
Db	1	GATGGTGACATGAGCCCTATCACTTAACGGCTGATATTCGATTTGTTCATGGTTTTT	60
Qy	21	LeuAsnPheAsnPheTyrThrLeuMetAsnIlyProIlyIysValAlaIleLeuThrAla	40
Db	61	CTTAACCTTAACTCTACACGGCTCATGAAACAACCTAATAAAAGTTGCATATGACAGCA	120
Qy	41	GIYGIYLeuAlaProCysLeuAsnSerAlaIleGIYSerLeuIleGIUArgTyrThrGIU	60
Db	121	GGCGGCTTGGCGGCTTGTGTAATTCGCCAATCGGTAGTTGATGCAACGTATACCGAA	180
Qy	61	IleAsnProSerIleGIUleIleCysTyrArgGIYGIYTYTYLYAGIYLeuLeuGIY	80
Db	181	ATCATACCTACATAGAAATATTTGCTATCGCGGGTATTAAGGCTGTGCTGGGC	240
Qy	81	AspSerTyrProValThrAlaGIUValArgIlyIysAlaGIYValIleuGIUArgPheGIY	100
Db	241	GATTTCTTATCCAGTAAAGCGCGAAGTGCTAAAGCGGGGTTCTGCAACGTTTGGC	300
Qy	101	GIYSerValIleGIYAsnSerArgValIlyLeuThrAsnValIlyAsnProCysValIlyAla	120
Db	301	GGTTCGTGATCGGCACACGCCGGCTCAAAATTGCACCAATGTCCAAAGACTGCGTGAACGC	360
Qy	121	GIYLeuValIlySGIUGIYGIUAsnProIlnIysValAlaIleAsnGIUleuValIlyAsn	140
Db	361	GGTTTGGCAAGAGGGGTGAAGATCCGCAAAAAGTCGGGCTGATCAATTTGGTTAAGAT	420
Qy	141	GIYValAspIleLeuHisThrIleGIYGIYAspAspThrAsnThrAlaAlaIleAsnLeu	160
Db	421	GGTGTCCATATTCGACACACCATCGCGCGCATGATACCAATCCGCGACACGCAATTTG	480
Qy	161	AlaAlaPheLeuAlaArgAsnAsnTyrGIYLeuThrValIleGIYLeuProIlyThrVal	180
Db	481	GCACACATTCCTGGCGCAAAATTAATTAACGCACTGACCGTCATGTGTTACTTAACCCGTC	540
Qy	181	AspAsnAspValIlePheProIleIlySGIHiserLeuGIYATrPThrAlaAlaGIUingIy	200
Db	541	GATTAACGACGTAATTCGATCAAGCAATCACTTAAGGTGTGACATCGCCGACGCAAGGC	600
Qy	201	AlaArgTyrPheMetAsnValValAlaGIUAsnAsnAlaAsnProArgMetLeuIleVal	220
Db	601	GGCGGTTATTTCAATGACAGCTGTGTGGCCGAAACCAACGCCACCAACGATCTGATCGTA	660
Qy	221	HisGIUAlaMetGIYArgAsnCysGIYTrpLeuThrAlaAlaThrAlaGIUingIyTyrArg	240
Db	661	CACGAGTGAAGGGGCGGTAACTGGGGGTGACCGGTGCACACCGCGCGACGAATATCCG	720
Qy	241	IlyIleLeuAsnAspArgAlaGIUTrpLeuProGIUleuGIYLeuThrArgGIUserTyrGIU	260
Db	721	AAATTAATGACGCGGTGCGAGTGTGTGCGGAAATGGGTGTGAATCTGTGAATCTTAATGA	780
Qy	261	ValHisAlaValAlaPheValProGIUMerAlaIleAsnLeuGIUAlaGIUAlaIlyAsnLeu	280
Db	781	GTGACCGGGATTTCTGTCGGAAATGGCGATCGACTTGAAGCCGAAAGCCGACGCCCTG	840

QY 281 ArgGUValMetAspLysValAspCysValAsnIlePheValSerGluGlyAlaGlyVal 300
 DB 841 CCGGAAAGTATGAGCAAAAGTCGATCGATCATCTTGGTTCCGAAGGTGCGCGCTC 900
 QY 301 GUAAIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
 DB 901 GAAGCATATGTCGCGGAAATGCAAGGCGCAAGAGTCCGCGCATGCTTCCGCG 960
 QY 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
 DB 961 CACATCAAACTGGATCGGTCACCTTGGTAATGGTTGCGAGCAATTCGCGCAGATG 1020
 QY 341 IlegIValAGluLysThrLeuValGlnLysSerGlyTrpPheAlaArgAlaSerAlaSer 360
 DB 1021 ATAGGCGCGGAAAAAACCCTGGTACAAAATCGGATACCTTCCCGCTGCTTCTGCTTC 1080
 QY 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
 DB 1081 AACGTTGACGACATCGCTTGTGATCAATCGTCGCGCATGTCGAGTCCGCTTC 1140
 QY 381 ArgArgGluSerGlyValIlegIYHisAspGluAspAsnGlyAsnValIleuArgAlaIle 400
 DB 1141 CCGCGGAGTCTGGCTGATCGGTCAAGAGCAAGACACGCAACCTGTTGCGTGCATC 1200
 QY 401 GluPheProArgIleLysGlyGlyLysProPheAsnIleAspThrAspTrpPheAsnSer 420
 DB 1201 GAGTTTCCGCGCATCAAGGCGGCAACCGTTCAATATGACACCGACTGTTCAATAC 1260
 QY 421 MetLeuSerGluIlegIYglnProLysGlyLysValGluValSerHis 437
 DB 1261 ATGTTGAGCGAAATCGCGCAGCTTAAGCGGTAAAGTCGAAGTCAGCCAC 1311
 RESULT 5
 US-10-320-874-15
 / Sequence 15, Application US/10320874
 / Publication No. US20030138909A1
 / GENERAL INFORMATION:
 / APPLICANT: Kofeas, Matcheos
 / APPLICANT: Odem, J. Martin
 / APPLICANT: No. US20030138909A1ton, Kelley C.
 / APPLICANT: Ye, Rick
 / TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
 / FILE REFERENCE: CL1619 US NA
 / CURRENT APPLICATION NUMBER: US/10320,874
 / CURRENT FILING DATE: 2002-12-17
 / PRIOR APPLICATION NUMBER: US/09/934,901
 / PRIOR FILING DATE: 2001-08-22
 / PRIOR APPLICATION NUMBER: 60/229,906
 / PRIOR FILING DATE: September 1, 2000
 / NUMBER OF SEQ ID NOS: 20
 / SOFTWARE: Microsoft Office 97
 / SEQ ID NO 15
 / LENGTH: 1311
 / TYPE: DNA
 / ORGANISM: METHYLOMONAS SP.
 US-10-874-15
 Alignment Scores:
 Pred. No.: 4.19e-253 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 15 Gaps: 0
 US-10-701-200-6 (1-437) x US-10-320-874-15 (1-1311)
 QY 1 AsnValAlaThrTrpProGlyHisLeuThrAlaAspIleArgPheCysHisTrpPhePhe 20
 DB 1 GAGTGATGCAATGCGCTATCATCAACGCGTGAATATGATTTGATTCGATTCGTTT 60
 QY 21 LeuAsnPheAsnPheTrpThrLeuMetAsnLysProLysLeuValAlaIleuThrAla 40

DB 61 CTTAATCTTAACCTTCTACACGCTCATGAACAACCTAAAAAAAGTTGCATATCTGACAGA 120
 QY 41 GlyGlyLeuAlaProCysLeuAsnSerAlaIlegIYSerLeuIlegIYArgTrpTrpGlu 60
 DB 121 GCGGCTTGGCGCTTGTGTAATTCGCAATCGGTAAGTTATATCAACGTTATACCGAA 180
 QY 61 IleAspProSerIlegIYIleCysTrpArgGlyGlyTrpLysGlyLeuLeuLeuGly 80
 DB 181 ATCGATCTTAGATGAATGAATCATTTGCTATCGCGGGGTATTAAGGCTGTTGCTGGGC 240
 QY 81 AspSerTrpProValThrAlaGluValArgLysLysValAlaGlyValIleuGlnArgPheGly 100
 DB 241 GATTTTATCCAGTAAACCGCCCAAGTCCGTAAAGGCGGCGTCTTCGCAAGCTTTTGGC 300
 QY 101 GYSerValIlegIYAsnSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
 DB 301 GGTTCGTGATCGGCAACGCGCGTCAATATGACCAATGTCAAGACTGCGTAACCC 360
 QY 121 GlyLeuValLysGluGlyGluAspProGlnLysValAlaAlaAspGlnLeuValLysAsp 140
 DB 361 GGTTCGTGTAAGAGGTTGAAGATCCGCAAAAAGTCGCGCTGATCAATTTGGTTAAGAT 420
 QY 141 GlyValAspIleLeuHisThrIlegIYIAspAspThrAsnThrAlaAlaAspLeu 160
 DB 421 GGTTCGATATTCGACACACATCGCGCGCATGATCAACAAATACGCGACACGCGATTG 480
 QY 161 AlaAlaPheLeuAlaArgAsnAsnTrpGlyLeuThrValIlegIYLeuProLysThrVal 180
 DB 481 GAGGCAATTCCTGCGCAAAATATTAACGATGACCGCTATGTTTACTTAACCCGTC 540
 QY 181 AspAsnAspValPheProIleLysGlnSerLeuGlyAlaTrpThrAlaAlaGluGlnGly 200
 DB 541 GATTAAGACGATATTTCCGATCAAGCAATCACTAGTGCTTGGACTGCGCCGACGAGGC 600
 QY 201 AlaArgTrpPheMetAsnValAlaAlaGluAsnAsnAlaAsnProArgMetLeuIleVal 220
 DB 601 GCGCGTTATTTATCATGAACGTGTGCGCCGCAAAACCAACCCACGATGCTGATCGTA 660
 QY 221 HisGluValMetGlyArgAsnCysGlyTrpLeuThrAlaAlaThrAlaGlnGluTrpArg 240
 DB 661 CACGAAGTATGAGCGCGTAACCTGCGCTGCGTGCAGCCGCTGCAACCGCGAGGAATATCC 720
 QY 241 LysLeuLeuAspArgAlaGluTrpLeuProGluLeuGlyLeuThrArgGluSerTrpGlu 260
 DB 721 AAATTAATGACCGTCCGATCGATGTTGCCGGAATGGTTTACTGTGAATCTTAATGA 780
 QY 261 ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluAlaLysArgLeu 280
 DB 781 GTGACGCGGTATTCGTTCCGGAATGCGGATCGACTGAAGCCGAAGCCAAAGCCCTG 840
 QY 281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluGlyAlaGlyVal 300
 DB 841 CCGGAAGTATGAGCAAAATGCAATTCGTCACATCTTGTTCGAAAGTCCGCGCGCTC 900
 QY 301 GUAAIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
 DB 901 GAAGCATATGTCGCGGAAATGCAAGGCGCAAGAGTCCGCGCATGCTTCCGCG 960
 QY 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
 DB 961 CACATCAAACTGGATCGGTCACCTTGGTAATGGTTGCGAGCAATTCGCGCAGATG 1020
 QY 341 IlegIValAGluLysThrLeuValGlnLysSerGlyTrpPheAlaArgAlaSerAlaSer 360
 DB 1021 ATAGGCGCGGAAAAAACCCTGGTACAAAATCGGATACCTTCCCGCTGCTTCTGCTTC 1080
 QY 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
 DB 1081 AACGTTGACGACATCGCTTGTGATCAATCGTCGCGCATGTCGAGTCCGCTTC 1140
 QY 381 ArgArgGluSerGlyValIlegIYHisAspGluAspAsnGlyAsnValIleuArgAlaIle 400

Db 1141 CGCCGCACTCTGGCGTATCGGTACGACGAAACACCGCACTGTTGGTCGCATC 1200
 Qy 401 GIupheProArG1eLyegLyLyPProPheasn11easPThrAsPTrPheasnser 420
 Db 1201 GAGTTTCGCGCATCAAGCGCGCAACCGTTCAATATGACACCGCATGTTCAATAGC 1260
 Qy 421 MetLeuSerGlu11eG1yGlnProLyseG1yLyLyValGluValSerHis 437
 Db 1261 ATGTTGAGCAAAATCGCGCACTAAAGCGGTAAAGTGAAGTCAAGCCAC 1311
 RESULT 6
 US-10-363-567-1
 / Sequence 1, Application US/10363567
 / Publication No. US20040077068A1
 / GENERAL INFORMATION:
 / APPLICANT: E. I. du Pont de Nemours & Company
 / TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
 / FILE REFERENCE: CL1903 PCT
 / CURRENT APPLICATION NUMBER: US/10/363,567
 / PRIOR FILING DATE: 2003-09-04
 / PRIOR APPLICATION NUMBER: US 60/229858
 / PRIOR FILING DATE: 2000-09-01
 / PRIOR APPLICATION NUMBER: US 60/229907
 / NUMBER OF SEQ ID NOS: 60
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO 1
 / LENGTH: 1311
 / TYPE: DNA
 / ORGANISM: Methylobomonas 16a
 US-10-363-567-1

Alignment Scores:

Pred. No.:	4,19e-253	Length:	1311
Score:	2284.00	Matches:	437
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	17	Gaps:	0

US-10-701-200-6 (1-437) x US-10-363-567-1 (1-1311)

Qy 1 AsPValThTrPProG1yH1eLeuThra1aAsP1aArGPhCyH1sTrPPhenPhe 20
 Db 1 GATGGGTGACATGCGCTTATCATCTTAACGGCTGATTCGATTTGTCATTTGTTT 60
 Qy 21 LeuAenPheAsnPhetThrLeuMetAsnLyPProLyseValAla11eLeuThra1a 40
 Db 61 CTTAACCTTAACTTCTTACACGCTCATGAACAACTTAAAGATTGCAATACTGACAGCA 120
 Qy 41 G1yG1yLeuAlaProCyLeuAenSerAla11eG1ySerLeu11eG1uArG1yThrg1u 60
 Db 121 GCGCGGCTTGGCGCTTGTGAATTCGCAATCGGAGTTGATCGAAGCTTATACGGA 180
 Qy 61 11eAsPProSer11eG1u11e11eCyS1y1yArG1yG1y1y1y1y1y1y1y1y1y1y 80
 Db 181 ATCGATCTTACGATGAAATCAATTTGCTATCGCGGCGGTAAAGCCCTGTTGCTGGGC 240
 Qy 81 AsPserTyProValThra1aG1uValArG1yLyValAla11eG1u1y1y1y1y1y1y1y 100
 Db 241 GATTTCTTATCCAGTAAACGCGCGAAAGTGGTAAAGGCGGAGTTTGTGCAAGCTTTGGC 300
 Qy 101 G1ySerVal11eG1yAenSerArGVal1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y 120
 Db 301 GGTTCGTATCGGCAACGCGCGTCAATGTAACCAAGTCAAGATCGCGGAAAGCC 360
 Qy 121 G1yLeuVal1y 140
 Db 361 GGTTCGTATCGGCAACGCGCGTCAATGTAACCAAGTCAAGATCGCGGAAAGCC 420
 Qy 141 G1yValAsP11eLeu11eThr11eG1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y 160
 Db 421 GGTTCGTATTTCTGACACCATCGCGGAGATGATACCAATACGCGACAGCGGATTTTG 480

Qy 161 Ala1aPheLeuAlaArGAsnAsnTy1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y 180
 Db 481 GCAGCATCTCGCGCAGAAATATTTACCGCATGACCGCTCATTTGTTTCTTAAACCGTTC 540
 Qy 181 AsPAsnAsPValPhePro11eLyegLyLyPProPheasn11easPThrAsPTrPheasnser 200
 Db 541 GATTAACGAGTATTTCCATCAAGCAATCACTAGGTGCTGAGTTCGCGCGCAGCAAGCC 600
 Qy 201 AlaArG1y1yPheMetAsnVal1a1aG1uAenAsnAlaAenProArGMetLeu11eVal 220
 Db 601 GCGCGTATTTTCAATGAACGTGTGGCGCAAAAACAACGCCAACCGCATGCTGATTCGTA 660
 Qy 221 HisG1uVal1y 240
 Db 661 CACGAAGGATGAGCGCGCTGACCTGCGCGCTGACCGCTGCAACCGCTGCAACCGCAGGAATATGCGC 720
 Qy 241 LysLeuLeuAsPAsPValaG1u1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y 260
 Db 721 AAATTAATCTGAGACCGGCGCGAGTGTGCGGAAATTTGGTTTGACTCTGTAATCTTATGA 780
 Qy 261 ValHisAlaValPheValPProG1uMetAla11eAsPLeuG1uAla1y1y1y1y1y1y1y1y1y 280
 Db 781 GTGCACGCGGTATTCGTTCCGGAATATGCGCATCGCATGGAAGCCGAAGCCGCGCTG 840
 Qy 281 ArG1uVal1y 300
 Db 841 CGCGAAGTGAATGACAAAGTGCATTCGCTCAACATCTTCGTTCCGAAGGTGCGCGCTC 900
 Qy 301 G1uAla11eVal1a1aG1uMetG1uAla1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y 320
 Db 901 GAGCTATCGTGCAGAAATGACAGGCCCAAAGCCAGGAAGTGCAGCGCATGCTTCGCGC 960
 Qy 321 His11eLyLeuAsPAlaVala1y 340
 Db 961 CACATCAAACTGAGATGCGGTCAACCTGTAATGTTCCGCGACCAATTCGCGCAGATG 1020
 Qy 341 11eG1yAlaG1uLy1y 360
 Db 1021 ATAGCGCGGAAAAACCTGCTGACMAAAATCGGGATCTTCGCGCGCTTCCTTC 1080
 Qy 361 AsnValaAsPAsPMeArG1y 380
 Db 1081 AACGTTGACGACATGCGTTGATCAATCGTGCAGCCGCTTGGCGGTGAGTCCGCTTC 1140
 Qy 381 ArGArG1ySerG1yVal11eG1yHisAsP1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y 400
 Db 1141 CGCGCGAGTCTGGGTATCGGTACGACGAAAGCAACGCGCAACGTGTGTCGATC 1200
 Qy 401 GIupheProArG1eLyegLyLyPProPheasn11easPThrAsPTrPheasnser 420
 Db 1201 GAGTTTCGCGCATCAAGGCGCGCAACCGTTCAATATCGACCGCATGTTCAATAGC 1260
 Qy 421 MetLeuSerGlu11eG1yGlnProLyseG1yLyLyValGluValSerHis 437
 Db 1261 ATGTTGAGCAAAATCGCGCACTAAAGCGGTAAAGTGAAGTCAAGCCAC 1311
 RESULT 7
 US-10-321-210-15
 / Sequence 15, Application US/10321210
 / Publication No. US20040115657A1
 / GENERAL INFORMATION:
 / APPLICANT: Kofas, Matheos
 / APPLICANT: Odom, J. Martin
 / APPLICANT: Norton, Kelley C.
 / APPLICANT: Ye, Rick
 / TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN
 / FILE REFERENCE: CL1619 US NA
 / CURRENT APPLICATION NUMBER: US/10/321,210
 / PRIOR FILING DATE: 2002-12-17
 / PRIOR APPLICATION NUMBER: US/09/934,901
 / PRIOR FILING DATE: 2001-08-22
 / PRIOR APPLICATION NUMBER: 60/229,906

; PRIOR FILING DATE: September 1, 2000
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 15
 ; LENGTH: 1311
 ; TYPE: DNA
 ; ORGANISM: METHYLOMONAS SP.
 ; US-10-321-210-15

Alignment Scores:

Pred. No.:	4,19e-253	Length:	1311
Score:	2284.00	Matches:	437
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	18	Gaps:	0

US-10-701-200-6 (1-437) x US-10-321-210-15 (1-1311)

```

QY 1 AapVala1ThrTyrProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhe 20
DB 1 GATGGTGCACATGGCCCTATCATCAAGCGGTGATTCGATTTCATTCGTTT 60
QY 21 LeuAsnPheAsnPheTyrThrLeuMetAsnLysProLysLysValAlaIleLeuThrAla 40
DB 61 CTTAACCTTACTTCTTACACGCTCATGAACAACTTAAAGTTTCAATCTGACAGCA 120
QY 41 GlyLysLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluTrpTyrThrGlu 60
DB 121 GCGCGCTTGGCCCTTGTGTAATTCGCAATCGAGTTGATTCGAACGTTATACGAA 180
QY 61 IleAspProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGlyLeuLeuLeuGly 80
DB 181 ATCGATCCATAGCATAGAAATCATTTGCTATCGCGCGGTATTAAGCCGTTCGGGCG 240
QY 81 AapSerTyrProValThrAlaGluValArgLysValGlyValLeuGluTrpPheGly 100
DB 241 GATTTCTTATCCAGTAAACGCGCCGAAGCGCTTAAAGCGCGGTTCGCAACGTTTGGC 300
QY 101 GlySerValIleGlyValAsnSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
DB 301 GGTTCGTGATCGGCAACGCGCGCTCAATGTACCAATGTCAAGACTGCGGAAACGC 360
QY 121 GlyLeuValLysGlyLysGlyLysAspProGluLysValAlaIleAspGluLeuValLysAsp 140
DB 361 GGTTCGTGATCGGCAACGCGCGCTCAATGTCAAGACTGCGGAAACGC 420
QY 141 GlyValAspIleLeuHisThrIleGlyLysAspThrAsnThrAlaAlaAspLeu 160
DB 421 GGTTCGTGATTCGTCACACCATCGCGCGCATGATACCAATCGGCAACGCGGATTTG 480
QY 161 AlaAlaPheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLysThrVal 180
DB 481 GGAGCATTCCTCGCCGGAATAATTAACGACGACGATGCTTGTACCTTAAACCGTC 540
QY 181 AapAsnAspValPheProIleLysGlnSerLeuGlyAlaTrpThrAlaAlaGluGlnGly 200
DB 541 GATTAACGACGATTTTCGATCAAGCAATCACTAGTGTCTTGACCTCCCGCAAGCAGGC 600
QY 201 AlaArgTyrPheMetAsnValAlaIleGluAsnAsnAlaAsnProArgMetLeuIleVal 220
DB 601 GCGCGTATTCATGAACTGCGTGGCGCGGAACCAACGCAACGCAATGCTGATCGTA 660
QY 221 HisGluValMetGlyLysArgAsnCysGlyTyrPheLeuThrAlaAlaThrAlaGlnGluTyrArg 240
DB 661 CACGAAGTATGGCGCGTAACTGCGCGCTGACCGCTGCAACCGCGCAAGAAATCCG 720
QY 241 LysLeuLeuAspArgLysGluTrpLeuProGluLeuGlyLysThrArgLysSerTyrGlu 260
DB 721 AATTTACTGAGACCGCTCCGAGTGTGGCGGAATTTGGGTTTGACTGTGAATTTTGA 780
QY 261 ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluAlaLysArgLeu 280
  
```

```

DB 781 GTGCACGCGCTATTCGTTCCGGAATGCGGATCGACTCGAAGCCGAACCGACCTG 840
QY 281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluValAlaGlyVal 300
DB 841 CGGGAAGTATGGACAAAGTCGATTCGTCGCAACATCTTCGTTCCGAAGTGGCGGCTC 900
QY 301 GluAlaIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
DB 901 GAAGCATTCGTCGCGGAATGCAAGCCGAAGGCGCAAGAGTCCGCGCATGCTCGCG 960
QY 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
DB 961 CACATCAAACTGATGCGGTCAACCTGTTAAATGTTGCGGAGCAATTCGCGCAATG 1020
QY 341 IleGlyAlaGluLysThrLeuValGlnLysSerGlyTyrPheAlaArgLysSerAlaSer 360
DB 1021 ATAGCGCGCGGAATAAACTTGTATCAAAATCGGATATCTTGCCCGCTTCGTTCC 1080
QY 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
DB 1081 AACGTTGACGATGCGTTGATCAAAATGTCGCGCGCATTCGCGGTGAGTCCGCTTC 1140
QY 381 ArgArgGluSerGlyValIleGlyHisAspGluAspAsnGlyValAsnValLeuArgAlaIle 400
DB 1141 CCGCGGAGTGTGGCGGTGATCGGTCAAGCAAGCAACGCAACGATTCGTCGATC 1200
QY 401 GluPheProArgIleLysGlyGlyLysProPheAsnIleAspThrAspTrpPheAsnSer 420
DB 1201 GATTTCCCGCATCAAGGCGCGCAACCGTTCAATGCAACCGCATCGTTCAATAGC 1260
QY 421 MetLeuSerGluIleGlyGlnProLysGlyLysValGluValSerHis 437
DB 1261 ATGTTGACGGAATCGCGCAGCCTAAAGCGGTAAAGTGAAGTCAGCGAC 1311

RESULT 8
US-10-369-493-38035
; Sequence 38035, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38035
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
; US-10-369-493-38035

Alignment Scores:
Pred. No.: 4,51e-66 Length: 552
Score: 659.00 Matches: 125
Percent Similarity: 82.61% Conservative: 27
Best Local Similarity: 67.93% Mismatches: 32
Query Match: 28.85% Indels: 0
DB: 17 Gaps: 0

US-10-701-200-6 (1-437) x US-10-369-493-38035 (1-552)
QY 33 LysLysValAlaIleLeuThrAlaGlyLysLeuAlaProCysLeuAsnSerAlaIleGly 52
DB 1 CAGAAAGTGCATAGCTGACCGCGGCTGCGCCCTCTCTCTTCGCGCGGCGGT 60
QY 53 SerLeuIleGluArgTyrThrGluIleAspProSerIleGluIleIleCysTyrArgGly 72
  
```


Db	61	GGCCGATCGAGCGCTATATGCGACANTCGGCGCTCGGAATGCAATCGACTCATTCGCTTACCGCTCC	120
Qy	73	GIYTYrLyBgLYleuLeuLeuLgLYsaBpSerTYrProValThrAlaGluValArgLYbLYs	92
Db	121	GGTTATTCAGGGGGGTCTTCTGGCGCGCGCTATCGAATCATCCAGAGATATATGCGCGAAAG	180
Qy	93	AlaGlyValLeuGlnAlaArgPheGlyGlySerValIleGlyAsnSerArgValLYsLeuThr	112
Db	181	GGCGATCTACTTCATCGTATGTTACGGCGCGCTCGCGCATGGCAACGCGCGTCAAGCTTGACC	240
Qy	113	AsnValLYsAspCYsaBpValLYsaArgLYleuValLYsGluGlyGluAspProGlnLYsVal	132
Db	241	AATGCGCGCGAATTGGCGCCAAAGCGCGCGCTCGTCAAGAAAGGCGACAATCGTTGGCGGTC	300
Qy	133	AlaAlaAspGlnLeuValLYsaBpGlyValAspIleLeuHisThrIleGlyGlyAspAsp	152
Db	301	GCCGCTGAACGGCGCTGCGCGCTACAGCGCATCACCATCTCCATCTCTTCGCTGGCGACGAC	360
Qy	153	ThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuValArgAsnAsnTYrGlyLeuThr	172
Db	361	ACCAACACACACCGCGCGGACCTTGCGGCTATCTTCGCGGCCCAAGGATATGACCTCAC	420
Qy	173	ValIleGlyLeuProLYserThrValaAspAsnAspValPheProIleLYsGlnSerLeuGly	192
Db	421	GTCGTCGGCGCTGCCAAGACGGGTGATATGACGTGTGCGCATCAAGCAGCTGCTCGGC	480
Qy	193	AlaTYrThrAlaAlaGluGlnGlyAlaArgTYrPheMetAsnValValAlaGluAsnAsn	212
Db	481	GCCTGAGACGGCGCGGAAGTCGGTGTCTTCTTCGACATGTCTGACACGACGACGACG	540
Qy	213	AlaAsnProArg	216
Db	541	GGCGGACCGAAA	552

```

RESULT 9
US-10-369-493-31795
; Sequence 31795, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 31795
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: Thermobifida fusca
US-10-369-493-31795

```

Alignment Scores:	
Pred. No.:	5,696-25
Score:	366.50
Percent Similarity:	42.49%
Best Local Similarity:	26.46%
Query Match:	19.42%
DB:	17
	Gaps:
	14
	Indels:
	81
	Mismatches:
	145
	Conservative:
	63
	Matches:
	1050
	Length:
	5,696-25

US-10-701-200-6 (1-437) x US-10-369-493-31795 (1-1050)

```

QY      23  PheAsnPheTyrThrLeuMetAsnLysProLysLysValAlaIleLeuThrAlaGlyGly 42
      |||  |||||  |||  ::|||  ::|||  |||||  |||||
Db      10  TTCAGGAGTACACGCAATG-----CGAGTCGGGGGTACTGACCGGGGGTGGC 57

```

43 LeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluAlaGlyTrpThrGluIleAsp 62
Db 58 GAGTCCCGGGCGGTCAAGCGGTCACTCCGGTGGGTGTCGTAAAGGCGATCAAG---GAA 114
Qy 63 ProSerIleGluIleIleCysTrpArgGlyGlyTrpLeuGlyLeuLeuLeuGlyIleAspSer 82
Db 115 TACGGATACAGATTGCTGCGAATTCCTGACCGGTGGCGGCCCTTGAAAGCGACACC 174
Qy 83 TyrProValThrAlaGluValAlaGlyLeuValAlaGlyValIleGluIleAspGlyIleSer 102
Db 175 ATGCCCGGTGACATGAGAGCGGTCCGC-----GGCATCTCCCGCGC---GGCGGTACC 225
Qy 103 ValIleGlyAsnSerArgValIleLeuThrAsnValIleAspCysValIleLysArgGlyLeu 122
Db 226 ATCTCGGCTCTCTCCGCGACCACTCATGAAGTCGAGGGCGGTGTCAAGCGG----- 279
Qy 123 ValIleGlyGluIleAspProGlnIleValAlaAlaAspGlnLeuValIleLysAspGlyVal 142
Db 280 -----GTCAGAGCAACATGGCCCGCGTGGGGTG 309
Qy 143 AspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAlaAlaAlaAspLeuAlaAla 162
Db 310 GACGGCGTCTGTCGCAATCGGCGGTGAGACACCTTCGCGGTGCGCGCGCAGCTC----- 363
Qy 163 PheLeuAlaAspAsnAsnTrpGlyLeuThrValIleGlyLeuProLysThrValAspAsn 182
Db 364 -----CACGACCACCGGTGTCAACGTCGGTGGTGTGCTCCCAAGACCATCAAC 411
Qy 183 AspValPheProIleGlyGlnSerLeuGlyAlaTrpThrAlaAlaGluGlnGlyAlaArg 202
Db 412 GACCTTCACGCCACCGCATACACCTTCGGCTTCGACACTGCGGTCAACATCGGACCGAG 471
Qy 203 TyrPheMetAsnValValAlaGluAsnAsnAlaAsnProArgMetLeuIleValHisGlu 222
Db 472 GCCATCGACCGGCTGCACACACCGCGAGTGCACACCGTGCCCTGGTGCCTC---GAG 528
Qy 223 ValMetGlyAspAsnCysGlyTrpLeuThr-----AlaAlaThrAlaGln 237
Db 529 GTCATGGGCGCGCACGCGGTGATGGCTGCACGCGGCGATGGCGCGCGCGCGC--- 585
Qy 238 GluTrpArgIleValLeuAspArgAlaGluTrpLeuProGluIleGluIleuThrArgGlu 257
Db 585 ----- 585
Qy 258 SerTrpGluValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluAla 277
Db 586 -----ACGTATCTCTGATCCCGCGAGCGCGCATTCGACATCGACGAGTCTCTC 633
Qy 278 LysValGlyLeuAspGluValMetAspTrpValAspCysValAsnIlePheValSerGluGly 297
Db 634 GCGTCACTCGAAAGCGCGCTTCAGACCACTACGCGCGCATCATCTGTGTCCGCGAGGCG 693
Qy 298 AlaGlyValGluAlaIleValAlaGluMetGlnAlaGlyGluIleGluValProArgAsp 317
Db 694 GCCACCGCGAAG-----GAGGGCCAGCTCACCTTGCCACCGCGGAGACCGGAC 741
Qy 318 AlaPheGlyHisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPhe 337
Db 742 TCCTTCGGCGACGTCCGCTCGCGCGCATC-----GGCCAGCGCGCTC 783
Qy 338 AlaGlnMetIleGlyAlaGluLysThrLeuValGlnLysSer-----GlyTrpPhe 354
Db 784 GCCGAGAGATGAGAGGCGGACCGGCAAGAAAGCGCGCTCGGTGTCTCTGGCGCACGTG 843
Qy 355 AlaAspAlaSerHisSerAsnValAspAspMetArgLeuIleLysSerCysAlaAspLeu 374
Db 844 CAGCGCGCGGTACCGCGCTTCGCGCTTCGACCGAGTCTGAGCCACCGCGCTGGCTGCAC 903
Qy 375 AlaValGluCysAlaPheArgArgGlnSerGlyValIleGlyHisAspGluAspGly 394
Db 904 GCCATCACCGCGGCTC-----CACACACAGACCTTCGGC 936
Qy 395 AsnVal-----LeuArgAlaIleGluPheProArgIle 405


```

; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 6060
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Streptomyces avermectilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1026)
; US-10-156-761-6060

Alignment Scores:
Pred. No.: 1,52e-23 Length: 1026
Score: 294.00 Matches: 101
Percent Similarity: 43.83% Conservative: 66
Best Local Similarity: 26.51% Mismatches: 140
Query Match: 12.87% Indels: 74
DB: 15 Gaps: 17

US-10-701-200-6 (1-437) x US-10-156-761-6060 (1-1026)

QY 34 LysValAlaIleLeuThrAlaGlyLeuAlaProCybLeuAsnSerAlaIleGlySer 53
DB 4 CGGGTCGAGTACTACCGGAGCGGCGAGCTGCCCGGCTCAACGCCGCTCCGCGGC 63
QY 54 LeuIleGluArgTyrThrGluIleAspProSerIleGluIleCysTyrArgGly 73
DB 64 CTCGTCCTCGTAAGGGGTGAG--GAGTACGGCTACGACTTGTGCTCCGCGAGCGC 120
QY 74 TyrIysGlyLeuLeuLeuGlyAspSerTyrProValThrAlaGluValArgIysAla 93
DB 121 TGGCGGGGACCACTCGAAGCGACGCC-----GTACGGCTCGACATCCCGCGTGC 174
QY 94 GlyValIleGluIleArgPheGlyIleGlySerValIleGlyAsnSerArgValIysLeuThrAsn 113
DB 175 GGCATCTCGCCCCGCG--GGCGGACGATCTCGGCTCTCGCGACCAACCCCTCAAG 231
QY 114 ValIysAspCybValIysArgGlyLeuValIysGluGlyIleAspProGluIleValAla 133
DB 232 CTCGACGACGCGATCGCGCG-----ATCAAGAG----- 261
QY 134 AlaAspGluLeuValIysAspGlyValAspIleLeuHisThrIleGlyGlyAspAspThr 153
DB 262 -----AACCTCGCCCAAGAGAGAGGTCGACGGCTCATCGCATCGCGCGGAGAGCAG 315
QY 154 AsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaAlaArgAsnAsnTyrGlyLeuThrVal 173
DB 316 CTCGGCGTCGCGCAGCGCTCAC-----GACGAGTACCGCGCTGCGCGC 360
QY 174 IleGlyLeuProIleThrValAlaAsnAspValPheProIleIleGluSerLeuGlyAla 193
DB 361 GTCGGCGTCCTCGAAGACCTACCAACGACTGTGCGGACCGACGATCACTGTCGCTTC 420
QY 194 TrpThrAlaAlaGluGlnIleValAlaArgTyrPhe-----MetAsnValAlaIleGluAsn 211
DB 421 GACACGCGCGTCGCGATCGCAGACCGAGCATCGACCGCTCCACACACCGCGGATCC 480
QY 212 AsnAlaAsnProArgMetLeuIleValHisGluValMetGlyIleArgAsnCybGlyTyrPhe 231
DB 481 CACATGCGC-----GTCTCTGTCGAGGTGATGGCGCGGACCGCGGCTGATC 531
QY 232 ThrAlaAlaThrAlaGlnIleTyrArgIysLeuLeuAspArgAlaGluTyrPheProGlu 251
DB 532 GCCATCTCACTCG----- 543

```

```

QY 252 LeuGlyLeuThrArgGluSerTyrGluValHisAlaValPheValProGluMetAlaIle 271
DB 544 --GGCTCGCGCGCGCGC-----AACGTCACTCATCTCCGAGCAGCGCTTC 591
QY 272 AspLeuGluAlaGluAlaValArgLeuArgGluValMetAspIysValAspCybValAsn 291
DB 592 GACGTGACCAAGGTGTGCGCTATGTACCTCGCGGTTCAAGGCTCGTACCGCGCATC 651
QY 292 IlePheValSerGluGlyAlaGlyValGluAlaIleValAlaGluMetGlnAlaIysGly 311
DB 652 GTGTGTCGCGCGAGGGCGC-----ATGCCAAGCAGCGC 687
QY 312 GlnGluValProArg-----AspAlaPheGlyHisIleLeuLeuAspAlaVal 327
DB 688 CAGATGTGCTTCAAGAGACAGATCTCCGATCTTCGCGCAGCTGCTTCGCGGCTC 747
QY 328 AsnProGlyIysTyrPheGlyGluGlnPheAlaGluMetIleGlyAlaGlu--LysThr 346
DB 748 -----GGCGAGTGTGCTGCCAAGAGATCGAAGAGCCACCGGACAGAGCCCGCAC 801
QY 347 LeuValGlnIysSerGlyTyrPheAlaArgAlaSerAlaSerAsnValAspAspMetArg 366
DB 802 ACGGTC-----CTCGGCGCATGTGACGCGCGGCGCAGCGCGCGTTCGACCGCTGG 855
QY 367 LeuIleIysSerCybAlaAspLeuAlaValGluCybAlaPheArgArgGluSerGlyVal 386
DB 856 CTCGCCACCCGCTTCGGCTCGACGCGCATCGAGGCGCTC----- 894
QY 387 IleGlyHisAspGluAspAsnGlyAsnVal-----LeuArgAlaIleGluPheProArg 404
DB 895 -----CGCGACGGCGACTTGGCGCAAGTGTGCTGCGCGGCGCAGCATCTGCTGT 948
QY 405 Ile 405
DB 949 GTC 951

RESULT 12
US-10-425-114-9204
; Sequence 9204, Application US/10425114
; Publication No. US2004003488B1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9204
; LENGTH: 1570
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700837684_FLI
; US-10-425-114-9204

Alignment Scores:
Pred. No.: 3.71e-23 Length: 1570
Score: 293.00 Matches: 113
Percent Similarity: 41.23% Conservative: 61
Best Local Similarity: 26.78% Mismatches: 174
Query Match: 12.83% Indels: 74
DB: 17 Gaps: 15

US-10-701-200-6 (1-437) x US-10-425-114-9204 (1-1570)

QY 34 LysValAlaIleLeuThrAlaGlyLeuAlaProCybLeuAsnSerAlaIleGlySer 53

```

```

Db      224 CAGGCTGATTTGTTACTTGTGGGGTCTGTCTCCGGGCTCAACCTCTCATTTAGGAA 283
Qy      54 Leuilegluargtyrthrleuprproserilegluilecysyrhrarglygly 73
Db      284 TTAGTGTGTGGCTTACACCATATGTATGGGGTGAAGAAAGTTCTTGAATCAATGAGAA 343
Qy      74 Tyrlyeglyleuleuleuglyaspseryrprovalthrleagluvalarglylyvala 93
Db      344 TATAGGGGTTTCTATGCTCGCATATACATCACTTAAACCCCT-----AAAAGTGTG 394
Qy      94 Glyvalleugluarphleglyglyservalileglyanaserthrvalylyseuthraen 113
Db      395 AATGATATACATTAAGGCTGGGAACTGTCTCGAATCACTACGA-----439
Qy      114 Vallyabpvcyvallyabpvcyvallyabpvcyvallyabpvcyvallyabpvcyval 133
Db      440 -----GGTGCACATGACCAAAAGATA--463
Qy      134 Alaaprglnleuvallyabpvcyvallyabpvcyvallyabpvcyvallyabpvcyval 153
Db      464 GTTGACAGATATCAAGATCGGGAAATCAATCAAGTTATATATGAGAGATGGAAGT 523
Qy      154 AenThrAlaAlaAlaAbpLeuAlaAlaPheLeuAlaArgAbpAbnTyrclyLeuThrVal 173
Db      524 CAGAAAGGTGCATCTGCATTTTGTGAGAGAGTCAAGAACGTGTCTCAAGTTTCAGTT 583
Qy      174 Ileglyleuprprolyethrvalleuprprolyethrvalleuprprolyethrvalle 193
Db      584 GTAGGATATCCCAAAACATAGATATGATATGATTCCTTATGATATGATTCCTTATGAT 643
Qy      194 TyrThrAlaAlaAlaGlucylAlaAlaArgTyrcPheMetAbnValAlaAlaAlaAla 213
Db      644 GACACTGCTGTGAGAGGCTCAACAGCATATTAATGACGACATCTTGAAGTGAAGT 703
Qy      214 AenProArgMetLeuileValHilegluValMetGlyArgAbnCyglyTyrclyLeuThrAla 233
Db      704 GTAGAAATGCGATAGGTGTGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 760
Qy      234 AlaThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 253
Db      761 -----ATGTAT 766
Qy      254 LeuThrArgGluSerTyrclyValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 273
Db      767 GCTACCTTGGCAGAGTGAACGTGATGTGTCTTATTCAGAGTCAACCTTTTACCTT 826
Qy      274 GluAlaAlaAla-----LysArgLeuArgGluValMetAbpLys 286
Db      827 GAAGGCTCGTGGAGCTTATGATATATAGAGAAAGACTAAAGAAATGGGCACATG 886
Qy      287 ValAbpCysValAbnAlaPheValSerGluAlaAlaAlaAlaAlaAlaAlaAlaAla 306
Db      887 GTT-----ATTGTTATGCTGAAGGACGACAGAACCTTCTTCTGAGAGT 934
Qy      307 MetGlnAlaLeuGlyGlnGluValProArgAbpAlaPheGlyHilegluValleuSerAla 326
Db      935 GTGCACATCATGACCAACAA-----GATGCTTCTGGAACCAAGCTTTTTCAGAT 985
Qy      327 ValAbnProGlyLysTyrcPheGlyGlnGlnPheAlaAlaMetileGlyAlaGluTyrc 346
Db      986 GTT-----GGCCTATGATATCCCAAAAGATTAGGATCATTTTCTGCACAGAGACA 1039
Qy      347 Leu-----ValGlnLysSerGlyTyrcPheAlaAlaArgLysSerAlaSerAla 361
Db      1040 CTGCCCATATCTCTCAATATACATAGATCCAACTATATATATCCGACCTTATCCAAACCAT 1099
Qy      362 ValAbpArgMetArgLeuileLysSerCysAlaAbpLeuAlaValGluCysAlaPheArg 381
Db      1100 GGCTGTGACAGATGATGACGACATCTTCTGCAAAAGTGCAGTTATGAGAGAGAAAG--- 1156
Qy      382 ArgGluSerGlyValileGlyHileAbpGluAbp-----AbnGlyAbnValleuArgAla 399

```

```

Db      1157 -----GACGTTTACATCGCTATACAACTGACCTTGTGAATGAGAA---CAAACTTAT 1207
Qy      400 IlegluPheProArgIleLysGlyLysProPheAbnIleAbpThrAsp-----Trp 417
Db      1208 ATACCTTCTATAGATATCACTGAGAGACAGAACGATGATATATATGATATGATGATG 1267
Qy      418 PheAbnSerMetLeuSerGluileGlyGlnPro-----LysGlyGlyLys 432
Db      1268 GCTAGG---CTTTTATCTTCAACAAATCAACCCAGCTTTTGTGATCCAAAGGTGACAAAT 1324
Qy      433 ValGlu 434
Db      1325 GAAGAG 1330

RESULT 13
US-10-425-114-6293
/ Sequence 6293, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 6293
/ LENGTH: 1765
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700559917_FLI
US-10-425-114-6293

Alignment Scores:
Pred. No.: 4,41e-23 Length: 1765
Score: 293.00 Matches: 113
Percent Similarity: 41.23% Conservative: 61
Bee Local Similarity: 26.78% Mismatches: 174
Query Match: 12.83% Indels: 74
DB: 17 Gaps: 15

US-10-701-200-6 (1-437) x US-10-425-114-6293 (1-1765)
Qy      34 LysValAlaAlaIleuThrAlaGlyLysLeuAlaProCysLeuAbnSerAlaIleGlySer 53
Db      371 CAGGCTGCTATTTGTTACTTGTGGGGTCTGTCTCCGGGCTCAACCTCTCATTTAGGAA 430
Qy      54 Leuilegluargtyrthrleuprproserilegluilecysyrhrarglygly 73
Db      431 TTAGTGTGTGGCTTACACCATATGTATGGGGTGAAGAAAGTTCTTGAATCAATGAGAA 490
Qy      74 Tyrlyeglyleuleuleuglyaspseryrprovalthrleagluvalarglylyvala 93
Db      491 TATAGGGGTTTCTATGCTCGCATATACATCACTTAAACCCCT-----AAAAGTGTG 541
Qy      94 Glyvalleugluarphleglyglyservalileglyanaserthrvalylyseuthraen 113
Db      542 AATGATATACATTAAGGCTGGGAACTGTCTCGAATCACTACGA-----456
Qy      114 Vallyabpvcyvallyabpvcyvallyabpvcyvallyabpvcyvallyabpvcyval 133
Db      587 -----GGTGCACATGACCAAAAGATA--610
Qy      134 Alaaprglnleuvallyabpvcyvallyabpvcyvallyabpvcyvallyabpvcyval 153
Db      611 GTTGACAGATATCAAGATCGGGAAATCAATCAAGTTATATATGAGAGATGGAAGT 670

```


QY 274 GIUAlaGIUAla-----LysArgLeuArgGIUAlaMetAspLys 286
 DB 1166 GAAGGCTGCTGAGACTATGAATATATAGAGAAAAGACTAAAGAAAATGGGCACATG 1225
 QY 287 ValAspCysValAsnIlePheValSerGIUAlaGIUAlaIleValAlaGIU 306
 DB 1226 GTT-----ATTGTTATTCCTGAAGGACAGACAGAACTTTGCTGAGAGT 1273
 QY 307 MetGlnAlaLysGIUAlaGIUAlaProArgAspAlaPheGlyHisIleLysLeuAspAla 326
 DB 1274 GTGCACTCCATGAGCAAAACA-----GATGCTTCTGGAACAAGCTTTTCAAGAT 1324
 QY 327 ValAsnProGlyLysTrpPheGIUAlaGIUAlaIleGlyValAlaGIUAla 346
 DB 1325 GTT-----GGCTATGATGATATCCCAAAAGATTAGGATCATTTTCTGCACGAAGACA 1378
 QY 347 Leu-----ValGlnLysSerGlyTrpPheAlaArgAlaSerAlaSerAsn 361
 DB 1379 CTGCCCATTACTCTCAAAATACATAGATCCAACTTATGATCCGACCTATTCCAAACAAAT 1438
 QY 362 ValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArg 381
 DB 1439 GCCTCTGACACGCTGACTGACACACTTTCTTCAAAAGTGCAGTTCATGAGCAATG--- 1495
 QY 382 ArgGluSerGlyValIleGlyHisAspGluAsp-----AsnGlyAsnValLeuArgAla 399
 DB 1496 -----GCAGGTACACTGGCTATACAGTGCAGCTTGTAATGAGA---CAACTTAT 1546
 QY 400 IleGluPheProArgIleLysGlyLysProPheAsnIleAspThrAsp-----Trp 417
 DB 1547 ATACCTTCTATAGATCACTGAGACAGAACCAAGTGTATGATGATGATGATGATGATGATG 1606
 QY 418 PheAsnSerMetLeuSerGluIleGlyGlnPro-----LysGlyGlyLys 432
 DB 1607 GCTAGG---CTTTATCTTCACAAATCAACCAAGCTTTTGGATGCCAAGGGTGAACAT 1663
 QY 433 ValGlu 434
 DB 1664 GAAGAG 1669
 RESULT 15
 US-10-424-599-141147
 / Sequence 141147, Application US/10424599
 / Publication No. US20040031072A1
 / GENERAL INFORMATION:
 / APPLICANT: La Rosa Thomas J
 / APPLICANT: Kovalic David K
 / APPLICANT: Zhou Yihua
 / APPLICANT: Cao Yongwei
 / TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 / FILE REFERENCE: 38-21(53223)B
 / CURRENT APPLICATION NUMBER: US/10/424, 599
 / NUMBER OF SEQ ID NOS: 285684
 / SEQ ID NO 141147
 / LENGTH: 4413
 / TYPE: DNA
 / ORGANISM: Glycine max
 / FEATURE:
 / OTHER INFORMATION: Clone ID: PAT_MRT3847_98468C.1
 US-10-424-599-141147
 Alignment Scores:
 Pted. No.: 1,69e-22 Length: 4413
 Score: 293.00 Matches: 113
 Percent Similarity: 41.23% Conservative: 61
 Best Local Similarity: 26.78% Mismatches: 174
 Query Match: 12.83% Indels: 74
 DB: 17 Gaps: 15
 US-10-701-200-6 (1-437) x US-10-424-599-141147 (1-4413)

QY 34 LysValAlaIleLeuThrAlaGlyLysLeuAlaProCysLeuAsnSerAlaIleGlySer 53
 DB 2954 CAGGCTGATATTGTTACTTGTGGGGCTCTGTGCTCGGGCTCAACATGATTAAGGAA 3013
 QY 54 LeuIleGluArgTrpThrGluIleAspProSerIleGluIleIleCysTrpArgGIU 73
 DB 3014 TTAGTGTGGCTTACACCATATGTTGGGGTGAAGAAAGTTCTTGAAATCAATGAGGA 3073
 QY 74 TrpLysGlyLeuLeuLeuGlyAspSerTrpProValThrAlaGluValArgLysVala 93
 DB 3074 TATAGGGGTTTCTATGCTGCAATACATCACTTAACCCCT-----AAAAGTGTG 3124
 QY 94 GlyValLeuGlnArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThrAsn 113
 DB 3125 AATGATATACATATAGGCTGGGGAACTGTCTCGAAATCAACGA----- 3169
 QY 114 ValLysAspCysValLysArgGlyLeuValLysGluGlyGluAspProGluLysVala 133
 DB 3170 -----GGTGGACATGACACCAAAAAGATA--- 3193
 QY 134 AlaAspGlnLeuValLysAspGlyValAspIleLeuHisThrIleGlyLysAspAspThr 153
 DB 3194 GTTGACAGTATTCAAGATCGGGAAATCAATCAAGTTTATTAATTTGAGAGAGTGAAC 3253
 QY 154 AsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTrpGlyLeuThrVal 173
 DB 3254 CAGAAAGGTGCATCTGCATTTTGGAGAACTGAGAAAGCTGCTCAAAAGTTTCAGTT 3313
 QY 174 IleGlyLeuProLysThrValAspAsnAspValPheProIleLysGlnSerLeuGlyAla 193
 DB 3314 GTAGGTATCCCAAAACCATATGATATGATATCCAGTTATGATGATGATGATGATGATG 3373
 QY 194 TrpThrAlaAlaGluGlnGlyAlaArgTrpPheMetAsnValValaGluAsnAsnAla 213
 DB 3374 GACACTGCTGTGGAGAGCTCAACGCTATTAATATGACGACATGTTGAAGTGAAGT 3433
 QY 214 AsnProArgMetLeuIleValHisGluValMetGlyArgAsnCysGlyTrpLeuThrAla 233
 DB 3434 GTAGAAATGCGCATGAGTGTTCATGATGATGATGATGATGATGATGATGATGATGATG 3490
 QY 234 AlaThrAlaGlnGluTrpArgLysLeuAspArgAlaGluTrpLeuProGluLeuGly 253
 DB 3491 -----ATGTAT 3496
 QY 254 LeuThrArgLysSerTrpGluValHisAlaValPheValProGluMetAlaIleAspLeu 273
 DB 3497 GCTACTCTTGCAGTGCAGAGCTGAGAGCTGATGTTGCTTAATTCAGAGTCACCCCTTTACCTT 3556
 QY 274 GIUAlaGIUAla-----LysArgLeuArgGIUAlaMetAspLys 286
 DB 3557 GAAGGCTGCTGAGACTATGAATATATAGAGAAAAGACTAAAGAAAATGGGCACATG 3616
 QY 287 ValAspCysValAsnIlePheValSerGIUAlaGIUAlaIleValAlaGIU 306
 DB 3617 GTT-----ATTGTTATTCCTGAAGGACAGACAGAACTTTGCTGAGAGT 3664
 QY 307 MetGlnAlaLysGIUAlaGIUAlaProArgAspAlaPheGlyHisIleLysLeuAspAla 326
 DB 3665 GTGCACTCCATGAGCAAAACA-----GATGCTTCTGGAACAAGCTTTTCAAGAT 3715
 QY 327 ValAsnProGlyLysTrpPheGIUAlaGIUAlaIleGlyValAlaGIUAla 346
 DB 3716 GTT-----GGCTATGATGATATCCCAAAAGATTAGGATCATTTTCTGCACGAAGACA 3769
 QY 347 Leu-----ValGlnLysSerGlyTrpPheAlaArgAlaSerAlaSerAsn 361
 DB 3770 CTGCCCATTACTCTCAAAATACATAGATCCAACTTATGATGCTGAGACTTATCCAAACAAAT 3829
 QY 362 ValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArg 381
 DB 3830 GCCTCTGACACGCTGACTGACACACTTTCTTCAAAAGTGCAGTTGATGAGCAATG--- 3886

```
Qy 382 ArgGluSerGlyValIleGlyHisAspGluAsp-----AsnGlyAsnValLeuArgAla 399
Db 3887 -----GCAGGTTACACTGGCTATCAAGTGGACTTGTGAATGAAGA---CAAACCTTAT 3937
Qy 400 IleGluPheProArgIleLysGlyLysProPheAsnIleAspThrAsp-----Trp 417
Db 3938 ATACCCCTTCTATAGAAATCACTGAGAGACAGAACCAACGTAAGTATTAAGTGAATGTGG 3997
Qy 418 PheAsnSerMetLeuSerGluIleGlyGlnPro-----LysGlyGlyLys 432
Db 3998 GCTAGG--CTTTATCTTCACAACAATCAACCCAGCTTTTGGATGCCAAGGATGACAAT 4054
Qy 433 ValGlu 434
Db 4055 GAAGAG 4060
```

Search completed: March 1, 2005, 21:18:32
JOB time : 4680 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2005, 14:11:02 / Search time 4373 Seconds
(without alignments)
3803.816 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVATPYPYHLADIRCFHWF.....FNSMLSEIQPKGKVEVSH 437

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q=/gen2.1/USFPO_spool/US10701200/runat.25022005.102832.10319/app.query.fasta.1.583
-DB=EST -QFMT=fastlap -SUFPTX=trc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-OCCALIGN=200 -THR SCORE=pcp -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10701200 @CNC 1.1 3437 @runat.25022005.102832.10319 -NCPU=6 -ICPU=3
-NO MAP -LARGEUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1341	58.7	1177	2	BE636682 rockefeller1
2	736	32.2	881	6	CD379187 PTMM04356
3	712	31.2	873	6	CD383878 PTMM09050
4	671	29.4	743	4	BM321461 rockefeller1
5	665.5	29.1	896	6	CD379057 PTMM04226
6	629	27.5	878	6	CD374942 PTMM00103
7	597.5	26.2	878	6	CD383585 PTMM08757
8	588	25.7	904	6	CD382389 PTMM07559
9	565.5	24.8	932	6	CD378198 PTMM03367

10	564	24.7	746	6	CD384626 PTMM09799
11	543.5	23.8	757	6	CD375650 PTMM00814
12	516.5	22.6	639	6	CD380358 PTMM05247
13	513.5	22.5	916	6	CD379504 PTMM04673
14	506	22.2	566	6	CD378314 PTMM03483
15	485.5	21.3	677	6	CD375439 PTMM00602
16	483.5	21.2	604	6	CD379552 PTMM04721
17	475.5	20.8	650	6	CD380700 PTMM05869
18	425	18.6	316	4	BM321460 rockefeller1
19	394.5	17.3	575	6	CD384340 PTMM09512
20	394.5	17.3	671	6	CD380085 PTMM05254
21	281.5	12.3	1797	3	CNS0A05M
22	278.5	12.2	1629	3	CL970486
23	268.5	11.8	303	4	BI307578
24	264.5	11.6	1734	3	CNS0A094
25	261.5	11.4	1596	9	CL957442
26	254	11.1	1398	3	CNS0A261
27	253	11.1	1737	3	CNS0A2H
28	251	11.0	990	9	CG155078
29	245.5	10.7	882	9	CG171162
30	244	10.7	982	7	CK278709
31	241	10.6	1036	9	CNS07402
32	235	10.3	701	6	CA208800
33	233	10.2	920	7	CV211012
34	231.5	10.1	764	4	BG365459
35	231.5	10.1	841	9	CNS0752F
36	231.5	10.1	1867	3	CNS0A305
37	230.5	10.1	628	6	CA157062
38	230.5	10.1	1661	3	CNS0A482
39	229	10.0	921	7	CV215682
40	228	10.0	1743	9	CL959679
41	227.5	10.0	802	6	CD849303
42	226.5	9.9	1430	3	CNS0A113
43	226	9.9	896	9	CG693232
44	226	9.9	897	7	CV209353
45	226	9.9	900	7	CV219943

ALIGNMENTS

RESULT 1
BE636682 LOCUS 1177 bp mRNA linear EST 03-JAN-2002
DEFINITION rockefeller.0.370 Mastigamoeba balamuthi lambda ZAP II library
Mastigamoeba balamuthi cDNA similar to pyrophosphate--fructose
6-phosphate 1-phosphotransferase, mRNA sequence.

ACCESSION BE636682
VERSION BE636682.2 GI:18055199

```

/strain="ATCC 30984"
/db xref="taxon:108607"
/clone_1.lib="Maatigameoba balamuthi lambda ZAP II Library"
/notes="Bryn: Phreatoecoba balamuthi"

```

Alignment Scores:

Pred. No.:	2,82e-152	Length:	25
Score:	1341.00	Matches:	11
Percent Similarity:	81.58%	Conservative:	52
Best Local Similarity:	67.38%	Mismatches:	70
Query Match:	58.71%	Indels:	0
DB:	2	Gaps:	0

US-10-701-200-6 (1-437) X BE636682 (1-1177)

QY	33	LYELVVALAIALELEuThrAlaGlyLYLeuAlaProCYLeuAsnSerAlaIGLY	52
Db	54	AMACRTGCGCCCTGCTCACTCCCGAGGCCCTCGCCCTCTCTCTCGCGCGTGGC	113
QY	53	SerLeuILEGLuArgTYrThrGLuILEaPProSerILEGLuILECYeTYrArgGLY	72
Db	114	GGCCTCATGAGGCGTACACAGAGATGTGCGCCGAGACAACATCATCTCTACCTCAAC	173
QY	73	GLTYrLYeGLYLeuLeuLeuGLYAspSerTYrProVALThrAlaGLuVALArgLYe	92
Db	174	GGGTTCACAGGGCTGCTGCTGGCGAGAAAGTGCTCTCGTACCCCGGCATAGGCTCAG	233
QY	93	AlaGLYValLEuGLuArgPheGLYGLYSerVALILEGLYAsnSerArgVALLYeLeuThr	112
Db	234	GGCAGCTCTGTGACACCGTCGGCGGAGCTGCATCGCAACTGGCGGTCAAGATGACC	293
QY	113	AsnVALLYeAspCYeVALLYeArgGLYLeuVALLYeGLuGLYGLuAspProGLuLYeVAL	132
Db	294	AAAGTGGCGGACTGCTCTCAAGCGCGGCTCTCGTCAAGAGGCGCAGAACCCAGCAAGTG	353
QY	133	AlaAlaAspGLInLeuVALLYeAspGLYValAspILEuAsnILEThrILEGLYLYAspAsp	152
Db	354	GGCGCGGACCAAGTGAATCAAGACCGCGTGCAGCGTCTGCACACGATGGCGGCGAGAC	413
QY	153	ThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTYrGLYLeuThr	172
Db	414	ACCAACAACGGCGGCGGAGCTGGCGGGTGCCTCAAGGGCGACCGGCTACACGCTCGC	473
QY	173	ValILEGLYLeuProLYeThrVALAspAsnAspVALPheProILEYGLInSerLEuGLY	192
Db	474	GTATGCGCTGCTGCCAAGACGATCGACAACGATTTGGCCCGTGGCGCAAGTCTGGGC	533
QY	193	AlATrPThrAlaAlaGLuGLInGLYAlaArgTYrPheMeAsnVALValaGLUAsnAsn	212
Db	534	GCCATGACGGCGCGAGCAGGCGTGGCTTCTTCAAGAACGTGTGCGCGAGCAGACG	593
QY	213	AlaAsnProArgMetLeuILEValHISGLUValMetGLYArgAsnCYeGLYTrPLeuThr	232
Db	594	GCCAAACCGCGCTGTGCTGTGCATCGACGAGTCAAGGGCCGCTGTGGGGTATCTTAGC	653
QY	233	AlaAlaThrAlaGLInGLYrTYrArgLYeLeuLeuAspArgAlaGLUTrPProGLInLeu	252
Db	654	GCCCAAGCGCGCGACACTACACCGCGCGAGCTCGCGCACCGGAGTTGGCCGCCAGCTC	713
QY	253	GLYLeuThrArgGLuSerTYrGLuVALHISAlaValPheVALProGLuMetAlaILEAsp	272
Db	714	GGCCACACCGCGAGCGCTACACATCCACGCGCTACGTGCGGAGATACCATGAC	773
QY	273	LeuGLuAlaGLuAlaLYeArgLeuArgGLUValMetAspLYeVALAspCYeVALAsnILE	292
Db	774	CTAAAGCGGAGCGCGCGCTGGCGCGCGCTCATAGAGCGGTGGCTGGTCAACATC	833
QY	293	PheVALSerGLuGLYAlaGLYAlaGLuAlaILEValaILEGLuMetGLuAlaLYeGLYGLIn	312
Db	834	TTCTCTCAGGAGGGCGCGGCGCATCAACGACATCGTGGGAGATGACCGCCAAAGGGGAG	893
QY	313	GLUVALProArgPheAlaPheGLYHISILEYLeuAsnAspAlaValAsnProGLYLYeTrP	332

Db	894	ACGGTCCCGCGCAGCCCTTCGGCGAAGTGAACCTCAACCCCGCGCGCTGG	953
Qy	333	PheGIyqIuGIInPheIaIaImeIIGeIyAlaGIuIyThrIeuValaGIuIySerGIy	352
Db	954	TTCCGCAAGCAAGTTTCGGCGGCATGTCGGCGCGGCAAGATCCTCTGTCAAGATCGGGC	1013
Qy	353	TyrPheAlaIArgIaSerIaSerIaSerIaValaAspAspMetArgIeuIleIySerCysAla	372
Db	1014	TACTCTCGGCGCTCGGCCCCCGGCACAGCGGAGACTCCGCTCATCAAGGCGATGGTG	1073
Qy	373	AspIeuAlaValaGIuIyCysAlaPheArgArgGIuSerGIyValaIIGeIyIaIaAspGIuAsp	392
Db	1074	GACTTCGCGCTCGACTGCGCGCTCGCGGCGAGGCTGCTGTATCGGCCACGACGAGGAG	1133
Qy	393	AsnGIyAsnAlaIeuArgIaIaIIGluPheProArgIleIyS	406
Db	1134	CGCAACGGCGTGGCGCGCGATCGATTTTGAGCGGCTCAAG	1175

RESULT 2

LOCUS	CD379187	881 bp	mRNA	linear	EST 31-MAY-2003
DEFINITION	P1YMN04356 Phaeodactylum tricornutum Uni-Zap XR phaeodactylum tricornutum cDNA 5', mRNA sequence.				

KEYWORDS

ORGANI

התקן

AUTHOR TITLE

JOURNAL
OF MEDICAL
EDUCATION

PUBMED
COMMENT

FEATURES

OPTICIN

Aj ianmen

Pred. No
Score:

Percent
Best Location

DB:

US-10-70

QY

Db

Qy

Qy 111 LeuThiRanValLysAspCyValIlyeArgGlyLeuValIysGluGlyLysProGln 130
Db 3 CTGCGAAACGTTGCCGACTGCATTAAAGAAAGGATACGTCACACGCGACGCCCTTG 62
Qy 131 LysValAlaAlaAspGlnLeuValIlyeAspGlyValAspIleLeuHieThrlIeGlyGly 150


```

Db      63 GAAGTGGCTTCGAGAGAGCTCATCAAGACCAAGTCCACTCCCTTCATACGATTGGCGGA
Qy      151 AAPAASPThrasnThrAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAenTyrgly 170
Db      123 GACGACACCAACACACAGAGTGTCTTACCTTTCCGACTCTGTGGAAAAGCAAAATGGC 182
Qy      171 ---LeuThrValIleGlyLeuProlySThrValAspAsnAspValPheProIlelysgln 189
Db      183 AAAGTCGGCGTATCGGTATGCCAAGACCATGACACAGATGTTATCTTATTTGATACAG 242
Qy      190 SerLeuGlyAlaTrpThrAlaAlaGluGlnGlyAlaArgTyPheMetAsnValValAla 209
Db      243 ACCTTCGGAGCCACACAGCTGCCGTGCGGGGGCGAGATTTTCGAGAACTTTCAT 302
Qy      210 GluAsnAspAlaAsnProArgMetLeuIleValHisGluValMetGlyArgAsnGly 229
Db      303 GAATGCACTGCCAATCTCGCATGCTTATTTCTCAGAGATGATGGACCGGATTCGGGA 362
Qy      230 TrpLeuThrAlaAlaThrAlaGlnGlyTyArgLyLeuLeuAspArgAlaGluTrpLeu 249
Db      363 TACCTTACGGCGCTACGGCAGAAAGTACCGGCATTTCTCAAT---GCCAAGACTTG 419
Qy      250 Pro-----GluLeuGlyLeuThrArgGluSer--TyrgluValHisAlaValPhe 265
Db      420 CCGCTCGGCTCGAGACTTCCCTTCACCGCAAGTCCCGCATTTATTCACGCTTTGG 479
Qy      266 ValProGluMetAlaIleAspLeuGluAlaGluValAlaValArgLeuArgGluValMetAsp 285
Db      480 ATTCGCCAATCTCAAGCTAGACTGTGTGTGCGAGCTCGCGGCTTTGAAGAAAGTCATGAG 539
Qy      286 LyValAspCyValAsnIlePheValSerGluGlyAlaGlyValAlaIleValAla 305
Db      540 GAAGTCGTTGGCTCAACATCTTTTCGGGAGGACCGGCTTCGACGAAATCGTCGCC 599
Qy      306 GluMetGlnAlaIleGlyGlnGluValProArgAspAlaPheGlyHisIleLySleuAsp 325
Db      600 GACATGAGAGCCCAACGCTGAGACCGGTGCGCGCATGCTTTGACACGTCATGTTGGCT 659
Qy      326 AlaValAsnProGlyLyserTrpPheGlyGluGlnPheAlaGluMetIleGlyAlaGluVal 345
Db      660 AAATATATCTCCGACAGATCTTCTCCCAACTTGCGGACAAATATGTGTCCGAAAGG 719
Qy      346 ThrLeuValGlnLyS--SerGlyTyrgPhe--AlaArgAlaSerAlaSerAsn--ValAspAsp 364
Db      720 ACCATTGTGCABAAGTTCGGGATACCTTTGGCCCGTTCGACGCGCCACGATTTTCGAT 779
Qy      365 MetArgLeuIle--LyserCyAlaAsp---LeuAlaValGluCyAlaPheArgArgG 383
Db      780 CGNCACTCATCGGGGCGCTGTGCCCAAGGAGTGGCGGTCCGCGCTTTTGACNGGA 839
Qy      383 u---SerGlyValIleGlyHisAspGluAspAsn 393
Db      840 CATTTCGGATGCTGTGGAAAGGATGAAAC 873

RESULT 3
LOCUS   CD383878                873 bp    mRNA    linear    EST 31-MAY-2003
DEFINITION   Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
ACCESSION   CD383878
VERSION     CD383878.1
KEYWORDS    GI:31259492
SOURCE      EST.
ORGANISM    Phaeodactylum tricornutum
            Phaeodactylum tricornutum
            Bacteria; Alveolates; Bacillariophyta; Bacillariophyceae;
            Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
            1 (bases 1 to 873)
REFERENCE   Scala S., Carls N., Falciatore A., Chiusano M.L. and Bowler C.
            Genome properties of the diatom Phaeodactylum tricornutum
            Plant Physiol. 129 (3), 993-1002 (2002)
JOURNAL    2211123
MEDLINE    22114555
PUBMED

```

```

COMMENT   Contact: Bowler C
            Laboratory of Molecular Plant Biology
            Stazione Zoologica 'Anton Dohrn'
            Villa Comunale I-80121, Napoli, Italy
            Tel: 39 081 583 3268/3211
            Fax: 39 081 764 1355
            Email: chris@alpha.szn.it
            Diatom EST Database (http://aveesthagen.sznbowler.com)
            Seq primer: T3 backward
            POLYA=Yes.

FEATURES
    source          Location/Qualifiers
    1..873
    /organism="Phaeodactylum tricornutum"
    /mol_type="mRNA"
    /db_xref="taxon:2850"
    /cell_line="CCMP632"
    /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
    /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
    Xho I"

ORIGIN
Alignment Scores:
Pred. No.:      1,18e-75      Length:      873
Score:          712.00        Matches:      151
Percent Similarity: 73.96%    Conservative: 45
Best Local Similarity: 56.98%  Mismatches:   60
Query Match:    31.17%        Indels:       9
DB:              6           Gaps:         4

US-10-701-200-6 (1-437) x CD383878 (1-873)

Qy      171 LeuThrValIleGlyLeuProlySThrValAspAsnAspValPheProIlelysglnSer 190
Db      7 GTCGCGCTCATCGGTATGCGCAAGACCATGACAGATGTTATCTTATTTGATACAGACC 66
Qy      191 LeuGlyAlaTrpThrAlaIleGluGlnGlyAlaArgTyPheMetAsnValValAlaGlu 210
Db      67 TTCGAGCGGACACACAGCTCCGCGAGGGGGGACGATTTTTCAGAAACGTTGCAATGAA 126
Qy      211 AsnAsnAlaAsnProArgMetLeuIleValHisGluValMetGlyTyrgAsnGlyTrp 220
Db      127 TGCACTGCCAATCTCTGATGCTTATCTTCACAGATGTATGGAGCGCATTCGGGATAC 186
Qy      231 LeuThrAlaAlaThrAlaGlnGlyTyArgLySleuLeuAspArgAlaGluTrpLeuPro 250
Db      187 CTTACGGCGGCTACGGCAAGAGTACCGGACATTTCTCAAT---GCCAAGACTTGGCC 243
Qy      251 -----GluLeuGlyLeuThrArgGluSer--TyrgluValHisAlaValPheVal 266
Db      244 GTCGCTCGGACTTGGCTTCCACCGCAAGTGTGCGGAGATTCACGCGCATTTGGATT 303
Qy      267 ProGluMetAlaIleAspLeuGluAlaGlyAlaValArgLeuArgGluValMetAspLyS 286
Db      304 CCGAATCTCAAGTCACTTGTGTCCGAGTCCGCGCTTTGAAGAAAGTCATGAGCGAA 363
Qy      287 ValAspCyValAsnIlePheValSerGluGlyAlaGlyValGluAlaIleValAlaGlu 306
Db      364 GTCGTTGGCTCAACATCTTTTTCGGCGAGGGGACCGGCTCGAGAAATCGTCGGGAC 423
Qy      424 ATCGAGCGCAACGGTGAGCCGCGCGCGCGCATGCTTTGACACGTCACGTTGGCTAAA 483
Qy      327 ValAsnProGlyLyserTrpPheGlyGluGlnPheAlaGluMetIleGlyAlaGluValTrp 346
Db      484 ATCAATCTCCGAGCACTTCTCCCAACTTGGCGGACAAATATTTGTCCGAAAAGACC 543
Qy      347 LeuValGlnLySerGlyTyrgPheAlaArgAlaSerAlaSerAsnValAspAspMetArg 366
Db      544 ATTGTCAAAAGTCGGATACCTTTGCCCTTCGCGCGCCGACCATTTTCATCCCAA 603
Qy      367 LeuIleLyserCyAlaAspLeuAlaValGluCyAlaPheArgArgGluSerGlyVal 386

```

Db 604 CTCATCGGGGCTGTGCGCCGANGCTGGCGCTGCCCGCTATTGACGACATTCGGGATGC 663

Qy 387 TleghYH1sAspG1uAspAenG1uAenVal1---Leu-ArgAla1leghYpHeProArg11 405

Db 664 ATGGAGACGAGTGAAGACAAACCAACGACGCCCATTTGAGCGATTGATTCAGTCGGCAT 723

Qy 405 e-Lyeg1Yg1YlypProPheAen11eAspTh1AspTrpPheAenSerMetLeuSer-Glu 424

Db 724 CAAGAGGNGGCAACCCITTGATTTCTCAGGAATGGGTCCNACAAATGCTCAAGGGA 783

Qy 425 TleghYg1n 427

Db 784 ATTGACAA 792

RESULT 4
BM321461
LOCUS
DEFINITION BM321461 743 bp mRNA linear EST 03-JAN-2002
rockefeller.0.1150 Mastigamoeba balamuthi lambda ZAP II library
Mastigamoeba balamuthi cDNA similar to pyrophosphate-fructose
6-phosphate 1-phosphotransferase (EC 2.7.1.9), mRNA sequence.

ACCESSION
VERSION BM321461.1 GI:18055867
KEYWORDS
SOURCE
ORGANISM Mastigamoeba balamuthi
Mastigamoeba balamuthi
Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE
AUTHORS 1 (bases 1 to 743)
Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
CONTACT: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockefeller.edu
Insert Length: 743 Std Error: 0.00
POLYA=yes

FEATURES
Source
1..743
/organism="Mastigamoeba balamuthi"
/mol_type="mRNA"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II library"
/note="syn: Phreatamoeba balamuthi"

ORIGIN

Alignment Scores:

Pred. No.:	9.42e-71	Length:	743
Score:	671.00	Matches:	126
Percent Similarity:	80.00%	Conservative:	30
Best Local Similarity:	64.62%	Mismatches:	39
Query Match:	29.38%	Indels:	1
DB:	4	Gaps:	0

US-10-701-200-6 (1-437) x BM321461 (1-743)

Qy 243 LeuAspAraG1uTrpLeuProG1uLeuG1yLeuTh1Arg1uSerTrpG1uValH1s 262

Db 1 CTCGGGCAACCGAGATTGGCGCCGAGCTCGGCGACAGCGGAGGCGTCAACGATCCAC 60

Qy 263 AlaValPheVal1ProG1uMetAla1leAspLeuG1uAlG1uAla1yAspArgLeuArg1u 282

Db 61 GCGGCTACGTCGCCGAGATGACATCGACTCAAGCGCGAGCGCCCGCTGCGCGCC 120

Qy 283 ValMetAspLeuVal1AspCysValAen11ePheValSerG1uG1yAlaG1yValG1uAla 302

Db 121 GTCAATGAGCGGTGGCTGCGTCAACATCTTCTCAGCGAGGCGCGGCGCATCAACGAC 180

Qy 303 TleValAlaG1uMetG1uAla1yG1yG1nG1uVal1ProA1gAspAlaPheG1yH1s1le 322

Db 181 ATCGCGCGGAGATGACCGCCAGGCGAGCGGCTCCGCGGACCC-TTCGGCCACGTC 239

Qy 323 LysLeuAspAlaVal1AspProG1yLysTrpPheG1yG1nG1uPheAlaG1uMet1leG1y 342

Db 240 AAGATCGACTTCATCAACCCCGCGCTGTTCCGACAGATTCGGCGGCGATGGTCGGC 299

Qy 343 AlaG1uYsTrpLeuVal1G1nLysSerG1yTrpPheAlaArgAlaSerAlaSerAenVal 362

Db 300 GCGGACAAAGTCTGTCGTCACAAAGTCGGGCTACTTCTCGCCCTGCGCCCGCCAAACGCC 359

Qy 363 AspAspMetArgLeu11eLysSerCysAlaAspLeuAlaValG1uCyAlaPheArgArg 382

Db 360 GAGGACCTCCCGCTCATCAAGGCGATGGTGAACCTGGCGCTGACATCGCGCGCTGCGCGC 419

Qy 383 GluSerG1yAla11eG1yH1sAspG1uAspAenG1uAenVal1leuAraG1a1leG1uPhe 402

Db 420 GAGGCTGCGCTGATCGGCGCACAGAGGCGCAACGCGCGTCTGCGCGCATCGAGTTT 479

Qy 403 ProArg1leYsG1yG1yLysProPheAen11eAspTh1AspTrpPheAenSerMetLeu 422

Db 480 GAGCGCTCAAGGCGGCGCAAGCGCTTCAACATCGACCAACCGTGGTTCACCCACTCTCTC 539

Qy 423 SerG1u1leG1yG1nProLysG1yG1yLysValG1uVal1SerH1s 437

Db 540 AACGAGATCGGCGCAGCCCAAGGCGCGCAAGGTCTCGTCCGCCAC 584

RESULT 5
CD379057
LOCUS
DEFINITION CD379057 896 bp mRNA linear EST 31-MAY-2003
PTM004226 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
tricornutum cDNA 5', mRNA sequence.

ACCESSION
VERSION CD379057
KEYWORDS
SOURCE
ORGANISM Phaeodactylum tricornutum
Phaeodactylum tricornutum
Eukaryota; Stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
REFERENCE
AUTHORS Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
Genome properties of the diatom Phaeodactylum tricornutum
Plant Physiol. 129 (3), 993-1002 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
CONTACT: Bowler C
Laboratory of Molecular Plant Biology
Stazione Zoologica 'Anton Dohrn'
Villa Comunale, I-80121, Napoli, Italy
Tel: 39 081 583 3268/3211
Fax: 39 081 764 1355
Email: chris@alpha.zn.dohrn.it
Diatom EST Database (http://aves.thagen.sznbowler.com)
Seq primer: T3 backward
POLYA=yes

FEATURES
Source
1..896
/organism="Phaeodactylum tricornutum"
/mol_type="mRNA"
/db_xref="taxon:2850"
/cell_line="CCMP632"
/clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
/note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
Xho I"

ORIGIN

Alignment Scores:

Pred. No.:	5.89e-70	Length:	896
Score:	665.50	Matches:	142
Percent Similarity:	73.83%	Conservative:	47
Best Local Similarity:	55.47%	Mismatches:	57

QY 391 GluAapanglyAanVal---LeuArlaAlaIleGluPheProArgIleGlyGly 409
 Db 599 GAAGGCAAAACCAACAGCCCATTCGAGGATTCAGTCGATCAAGGGTGCAAA 658
 QY 410 -ProheanleleagThraP-TpPheanSerMetLeuSer-GluIleGlyGln 427
 Db 659 CCCCTTGAATTTCTCANGAAATGGNTTCCAAATGCTCAAGGAAATTTGACAA 715

RESULT 7
 CD383585 878 bp mRNA linear EST 31-MAY-2003
 LOCUS PTM008757 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 DEFINITION tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD383585
 VERSION CD383585.1 GI:31259199
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 878)
 AUTHORS Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
 TITLE Genome properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 22111123
 PUBMED 12114555
 COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: christ@alpha.szn.it
 Dictionom EST Database (http://aveschagen.szbwler.com)
 Seg primer: T3 backward
 POLYA=yes

FEATURES
 source Location/Qualifiers
 1..878
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"

ORIGIN
 Alignment Scores: 1.15e-61 Length: 878
 Pred. No.: 597.50 Matches: 129
 Score: 70.894 Conservative: 39
 Percent Similarity: 54.43% Mismatches: 61
 Best Local Similarity: 26.16% Indels: 8
 Query Match: 6 Gaps: 5

US-10-701-200-6 (1-437) x CD383585 (1-878)

QY 70 TYRAGlyGlyTYrTyGlyLeuLeuLeuGlyAaPserTy-ProValThraIaGluVal 89
 Db 62 TACCTTGGGGGTACAAAGAAATGCTTACGGAGATTCATCGTCAATCCGGAATCA 121
 QY 90 ArgLybLybAlaGlyValleuGlnArgPheGlyGlySerValIleGlyAanSerArgVal 109
 Db 122 TGGGATTCACTCGACTCTCTGAATACGGTGGAGGATCGCCCATTTGGAATTCGCGCGTC 181
 QY 110 LyeLeuThraSerValLybAaPcybValLybAaGlyLeuValLybGluGlyGluAaPPro 129
 Db 183 AAGCTGACCAAGCGTGGCCATTCGATTAAAGAGGATACGTCACACGCGACGCGCC 241
 QY 130 GlnLybValAlaAlaAaPglInLeuValLybAaPGLyValAaPILeuhIsthriIleGly 149
 Db 242 TTAGAAGTACGCTTCACAGACGCTCATCAAGACCAAGTCCAGCTGCTTACGATTGGC 301

QY 150 GlyAaPaePThraAntThraAlaAlaAaPleuAlaAlaPheLeuAlaAaPAnAntTy 169
 Db 302 GGAAGCAGACCAACACACAAAGCTGACCTTTCCGACTTCGAGAAAGCAAT 361
 QY 170 Gly---LeuThraValIleGlyLeuProbyThraValAaPAnAaPValPheProIleLy 188
 Db 362 GGCAAGTCGCGCTCATCGTATGTCACCAAGACATCGACCAAGANTTTATCTTATTC 421
 QY 189 GlnSerLeuGlyAlaATPThraAlaIleGluGlnGlyAlaAaGTYrPheMechAanVal 208
 Db 422 CAGACCTTCGAGCGGACAGACAGCTGCCGTACAGGGGCGACCGCTTTTCAGAGAACGTTGTC 481
 QY 209 AlaGluAanAaAlaAaPProArgMetLeuIleValIleGluValAaMetGlyAaPAnGly 228
 Db 482 AATGAATGCACTGCCAAATCTCGCATGTCTTATTCACAGATGTATGAGACGCAATTCG 541
 QY 229 GlyTTPLeuThraAlaAlaThraAlaGlnGluTyArgLybLeuLeuAaPArgAlaGluTTP 248
 Db 542 GATTAACCTTACGGCGCGCTTACCGGACAAAGTACCGGCACATTTCAAC---GCCCAAGC 598
 QY 249 LeuPro-----GluLeuGlyLeuThraArgGlySer---TYrGluValIleAlaVal 264
 Db 599 TTGCCCGTCGCGCTCGACTTCGCCCTTCCACCGCAAGTCTGCCCGCATATTCACCGCAT 658
 QY 265 -PheValProGluMetAlaIleAaPLeuGluAlaGluAlaLybAaPArgGluValMe 284
 Db 659 TTGATTCGCCAATCCAACTGACTGCTGCGCGCAAGTGGCGCTTTGANAANAAGTCAT 718
 QY 284 LAsPlybValAaPcybValAanIle---PheValSerGluGlyAlaGly 299
 Db 719 GGGCGAATCGGTTCCGTTAACATCTTTTTCNGGAAAGGCAACCGGC 767

RESULT 8
 CD382389 904 bp mRNA linear EST 31-MAY-2003
 LOCUS PTM007559 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 DEFINITION tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD382389
 VERSION CD382389.1 GI:31258003
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 904)
 AUTHORS Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
 TITLE Genome properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 22111123
 PUBMED 12114555
 COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: christ@alpha.szn.it
 Dictionom EST Database (http://aveschagen.szbwler.com)
 Seg primer: T3 backward
 POLYA=yes

FEATURES
 source Location/Qualifiers
 1..904
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"

ORIGIN
 Alignment Scores: 1.74e-60 Length: 904
 Pred. No.: 174e-60

Score: 588.00 Matches: 131
 Percent Similarity: 71.84% Conservative: 45
 Best Local Similarity: 53.47% Mismatches: 59
 Query Match: 25.74% Indels: 10
 DB: 6 Gaps: 4

US-10-701-200-6 (1-437) x CD382389 (1-904)

QY 157 AAlaAAspLeuAlaAlaPheLeuAlaArgAsnAsnTyrGly---LeuThrValIleGly 175
 Db 2 GGTGCTAACCTTTCGCACTACCTGTGGAAACACAAATGGCAAAATCGCGCATCGGT 61
 QY 176 LeuProLyThrValAlaPheAsnAspValPheProIleLeuGlnSerLeuGlyAlaTyrThr 195
 Db 62 ATGCCCAAGACCATCCACAAACGATGTTTATCCATTGTCCAGACCTTCGAGCGGCACAA 121
 QY 196 AAlaAGlnGlnGlyAlaArgTyrPheMetAspValValAGlnLysAsnAlaAspPro 215
 Db 122 GCTGCCGTCAGAGGGGACGCTTTTCGAGAACGTTGTCANATGATGCACTGCCAAATCCT 181
 QY 216 ArgMetLeuIleValIleGlnValMetGlyArgAsnCyGlyTyrPheThrAlaAlaThr 235
 Db 182 CGATCTTATTTCTCCACGAGTGTATGGACGCGATTCGAGATCCTTACGCGCGCTAG 241
 QY 236 AAlaGlnGlyTyrArgLysLeuLeuAspArgAlaGlnTyrLeuPro-----GlnLeu 252
 Db 242 GCACAAAGATACCGCACATCTTCAAC---GCCCAAGCGTTCGCGCTCGGCTCGACTG 298
 QY 253 GlyLeuThrArgGlnSer---TyrGlnValHisAlaValAlaPheValProGlnMetAlaIle 271
 Db 299 CCTTCCACCGCAAGTCTGCCCGCATTCACGCAATTTGATTCGCCAATCCAGCTA 358
 QY 272 AspLeuGlnAlaGlnValLysArgLeuArgGlnValMetAspLysValAspCyValAsn 291
 Db 359 GACTGTGTCGCGAGTCGCGCGCTTGAAGAAAGTATGAGACCAATCCGCTCGCTCAAC 418
 QY 292 IlePheValSerGlnGlyAlaGlyValAGlnValIleValAlaGlnMetGlnAlaLysGly 311
 Db 419 ATCTTTTTCGCGCAGGCGACGCGCGCTGCANGAATGCTCGCGGACATGAGGCGCATGT 478
 QY 312 GlnGlnValProArgAspAlaPheGlyHisIleLysLeuAspAlaValAsnProGlyLys 331
 Db 479 GACGCGTCCGCGCGCATGCTTGAACAGTCACTGCTAAATCAATCCCGGACAG 538
 QY 332 TrpPheGlyGlnGlnPheAlaGlnMetIleGlyAlaGlnLysThrLeuValGlnLysSer 351
 Db 539 TTTCTTCTCCCAACACTTGGCGGACATATTGTCGCCGAAAGACATTTGTGCAAAAGTCG 598
 QY 352 GlyTyrPheAlaArg-AlaSerAlaSerAsn-ValAspAspMetArgLysIle-LysSer 370
 Db 599 GGATACCTTGGCGGTTCCGACGCGGCAACNATTTTNGATCGCCAACTCATCGCGGCGC 658
 QY 371 CysAla-AspLeuAlaValAGlnCysValAlaPheArgGlnSerGlyValIleGlyHisAs 390
 Db 659 TGTNCCGAGNGTCGCGTNGCGCGCTTATTTGACNAAATTCCTCCGNAATTCNTGGGA 718
 QY 390 pGluAsp 392
 Db 719 CAAGAT 725

RESULT 9
 CD378198 932 bp mRNA linear EST 31-May-2003
 LOCUS PTM003367 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 DEFINITION tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD378198
 VERSION CD378198.1 GI:31253812
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Bacteria; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 REFERENCE 1 (baaes 1 to 932)

AUTHORS Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
 TITLE Genome properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 22111123
 PUBMED 12114555

COMMENT
 Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Dicom EST Database (http://aves.thegen.sznbowl.com)
 Seq primer: T3 backward
 POLYA=Yes.

FEATURES

source location/Qualifiers
 1..932
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"

ORIGIN

Alignment Scores:

Pred. No.: 1,02e-57 Length: 932
 Score: 555.50 Matches: 115
 Percent Similarity: 77.27% Conservative: 38
 Best Local Similarity: 58.08% Mismatches: 39
 Query Match: 24.76% Indels: 6
 DB: 6 Gaps: 3

US-10-701-200-6 (1-437) x CD378198 (1-932)

QY 164 LeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLyThrValAspAsnAsp 183
 Db 31 TTGGAAGAACACAAATGGCAAAATCGCGCATCGGTATGCCAAGACATCGCAACGAT 90
 QY 184 ValPheProIleLysGlnSerLeuGlyAlaTyrThrAlaAlaGlnGlnGlyAlaArgTyr 203
 Db 91 GTTATTCCTATTGTACAGACTTCGAGCGGACACAGCTGCCGTCGACGAGGCGACATTT 150
 QY 204 PheMetAspValValAGlnLysAsnAlaAsnProArgMetLeuIleValHisGlnVal 223
 Db 151 TTGGAAGACGTTGTCATGATGCACTGCAATCTCGCATCTTATTTCTCCACGAGTGT 210
 QY 224 MetGlyArgAsnCyGlyTyrPheThrAlaAlaThrAlaGlnGlnLysArgLysLeu 243
 Db 211 ATGGGACGGGATTCGGGATACCTTACGGCGCTAGCGGACAAAGTACCGGACATTTCT 270
 QY 244 AspArgAlaGlnTyrLeuPro-----GlnLeuGlyLeuThrArgGlnSer---Tyr 259
 Db 271 AAT---GCCCAAGACTTGGCGCTCGCTCGACTTGGCCCTTCACCGCAAGTCTGCCGC 327
 QY 260 GlnValHisAlaValPheValProGlnMetAlaIleAspLeuGlnAlaGlnLysArg 279
 Db 328 GATATTTCACGCAATTTGATTCCTCCGAACTCAAGCTGATGATGTGCGAGTGGCGCGT 387
 QY 280 LeuArgGlnValMetAspLysValAspCyValAsnIlePheValSerGlnGlyAlaGly 299
 Db 388 TTGAAGAAAGTCATGAGACGAAGTCCGTTGCGTCAACATCTTTTGGCAGAGGACCGGCG 447
 QY 300 -ValGlnAlaIleValAlaGlnMetGlnAlaLysGlnGlnGlnValProArgAspAla 319
 Db 448 CGTGCAGGAATCGTCGCGACATGAGGCGCAACGNGAAGCGTGCCTCGCGGATGCCCTT 507
 QY 319 eGlyHisIleLysLeuAspAlaValAsnProGlyLysTyrPheGlyGlnGlnPheAla 339
 Db 508 TGGACAGTCACGTTGGCTAAATCAATCCGGAAGTATCTTCTCCCAACTTGGCGGA 567
 QY 339 mMetIleGlyAlaGlnLysThrLeuValGlnLysSerGlyTyrPheAlaArg 356

RESULT 10	746 bp	mRNA	linear	EST 31-MAY-2000
Db	CD384626	Phaeodactylum tricornutum	Uni-Zap XR	Phaeodactylum tricornutum cDNA 5', mRNA sequence.
LOCUS	CD384626			
DEFINITION				
ACCESSION	CD384626.1	GI:31260240		
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
COMMENT				
FEATURES				
Source				
Location/Qualifiers				
1..746				
/organism="Phaeodactylum tricornutum"				
/mol_type="mRNA"				
/db_xref="taxon:2850"				
/cell_line="CCMP632"				
/clone_lib="Phaeodactylum tricornutum Uni-Zap XR"				
/note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"				
Alignment Scores:				
Pred. No.:	1..11e-57	Length:	746	
Score:	564.00	Matches:	113	
Percent Similarity:	73.53%	Conservative:	37	
Best Local Similarity:	55.39%	Mismatches:	48	
Query Match:	24.69%	Indels:	6	
DB:	6	Gaps:	4	
US-10-701-200-6 (1-437) x CD384626 (1-746)				
Qy	229	GLYTPLeuthrAlaAlaThralagInguIuYrArGlySLeuleAspArgAlaGluTrp	248	
Db	9	GGATPACTTACGGCCCGCTACGGACAAAGTACCCGACATTCTCAAC---GCCCAAGGC	65	
Qy	249	LeuPro-----GluLeuGlyLeuThrArGJuseR--"TyrGluAlaHleAlaVal	264	
Db	66	TTGGCCCGCTCGGCTCGGACCTTGCCCTTCCACCGCAAGTCTGCCCGCGATATTCACGCATT	125	
Qy	265	PheValProGluMetAlaIleLeuPheGluAlaGluAlaAlaValArGLeuArgGluValMet	284	
Db	126	TGGATTCCCACTCAAGCTAGACTTGCTGGCCGCGAGTCCGCGCTTTGAGAGAAAGTATG	185	
Qy	285	AspIleValAspCysValAenIlePheValIserGluGlyAlaGlyValGluAlaIleVal	304	
Db	186	GACCAAGTCGGTTCGCTCAACATCTTTTCGGCGAGGCGACGGCGCTCAGAGAAATGTC	245	
Qy	305	AlaGluMetGlnAlaIleValArgGluGluValProArGAspAlaPheGlyHisIleLeuLeu	324	
Db	246	CGCGACATGGAAGCGCAATGATGACGCGCGCGCGGATGCTTTTGACACATGACGTTG	305	
Qy	325	AspAlaValAlaSerProGlyIleValTrpPheGluGluGluGluGluGluGluGluGluGlu	344	

[illegible]

```

Db      1 GGGGAGACGACACCAACACAGCTGCTACCTTCCGACTGCTGTGAAAAAGCAC 60
Qy      168 Aen-TyrglyLeuThrValIleGlyLeuProlyThrValAspAsnAspValPheProI 187
Db      61 MATGGCAAGATGCGCGTCACTGCTATGCGCAAGACATCGACACGATTTATCTTAT 120
Qy      187 elyGgInserLeuGlyValATrThrAlaAlaGInGInGlyAlaArgTy-PheMetAsnVa 207
Db      121 TGACACAGACCTTCGGAGCGGACACAGCTGCGGACAGGAGGCGACGATTTTCAGAACG 180
Qy      207 ValAlaGluAsnAsnAlaAsnProArgMetLeuIleValHisGluValMetGlyArgAs 227
Db      181 TGCAATGATGATGACATGCGCAATCTCGCATGCTTATTCACAGATGATAGGACGCA 240
Qy      227 ncyGgIyTrPLeuThrAlaAlaThrAlaGInGInGlyArgTyArgTyLeuAspArgAlaG 247
Db      241 TTCGGATACCTTACGGCGGCTACGGGACAGAAAGTACCGGACATTTCTCAAT--GCCA 297
Qy      247 uTrPLeuPro-----GluLeuGlyLeuThrArgGluSer---TyrglyValHisAl 263
Db      298 AGACTTGCCCGCTGCGCTCGGACTTGCCCTTCCACCGCAAGTGTGCCGCGGATTCACCG 357
Qy      263 eValPheValProGluMetAlaIleAspLeuGluAlaGluAlaLeuArgLeuArgGluVa 283
Db      358 CATTTGGATTCGCGAATCMAAGCTAGACTGTGCGCGAGTCCGCGCGCTTTGAAAGAAAG 417
Qy      283 lMetAspLyValAspCysValAsnAlaPheValSerGluGlyValAlaGlyValGluAlaI 303
Db      418 CATGACGAAAGTGGTGGTGGTCAACATCTTTTCCGCGGACAGGACGCGGATGANGAAAT 477
Qy      303 eValAlaGluMetGlnAlaLeuGlyGInGluValProArgAspAlaPheGlyHisIleLy 323
Db      478 CGTCCGGACATGAGAACGCAAGGTGAAACCGTCCGCGGAGTCCCTTGGACACCTCAC 537
Qy      323 eLeuAspAlaValAsn-ProGlyLyLeuTrPheGlyGluGInPhe---AlaGInMetIleG 342
Db      538 GTTGGCTMAATCAATNCCGGGACAGTACTCTTCCACACCTTNGCGGACATATTTGG 597
Qy      342 lValaGluTrPheLeu-ValGInLySerGlyTyPheAlaArgAlaSerAlaSerAs 361
Db      598 NGCCGAAAAACCATTTGNGCAGAAAGTGGGATCTTTGCGGATTCGCCGCGGACCA 657
Qy      361 n 361
Db      658 C 658

RESULT 12
CD380358      639 bp      mRNA      linear      EST 31-MAY-2003
LOCUS      PTM05527 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
DEFINITION      tricornutum cDNA 5', mRNA sequence.
ACCESSION      CD380358
VERSION      CD380358.1 GI:31255972
KEYWORDS      EST.
SOURCE      Phaeodactylum tricornutum
ORGANISM      Phaeodactylum tricornutum
Bukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
1 (bases 1 to 639)
Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
Genome properties of the diatom Phaeodactylum tricornutum
Plant Physiol. 129 (3), 993-1002 (2002)
2211123
12114555
Contact: Bowler C
Laboratory of Molecular Plant Biology
Stazione Zoologica 'Anton Dohrn'
Villa Comunale, I-80121, Napoli, Italy
Tel: 39 081 583 3268/3211
Fax: 39 081 764 1355
Email: chris@alpa.szn.it
Diatom EST Database (http://aveethagen.sznbowler.com)

```

```

FEATURES
source      Seg primer: T3 backward
            POLYA=No.
            Location/Qualifiers
            1..639
            /organism="Phaeodactylum tricornutum"
            /mol_type="mRNA"
            /db_xref="taxon:2850"
            /cell_line="CMP632"
            /clone_1ib="Phaeodactylum tricornutum Uni-Zap XR"
            /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
            Xho I"

ORIGIN
Alignment Scores:
Pred. No.:      5,54e-52      Length:      639
Score:          516.50      Matches:      96
Percent Similarity: 74.57%      Conservative: 33
Beet Local Similarity: 55.49%      Mismatches: 43
Query Match:    22.61%      Indels:      1
DB:             Gaps:        1

US-10-701-200-6 (1-437) x CD380358 (1-639)

Qy      256 ArgGluSerTyrglyValHisAlaValPheValProGluMetAlaIleAspLeuGluAla 275
Db      23 AAGTGTGCGCGGAGATTCACGCCATTTGGATTCGCACTCAAGCTAGACTGTGCGCC 82
Qy      276 GluAlaLeuArgLeuArgGluValMetAspLyValAspCysValAsnAlaPheValSer 295
Db      83 GAGTCCGCGCGGTTTGAAGAAAGTCATGACGAGTGGTGGTCAACATCTTTTCCGCC 142
Qy      296 GluGlyAlaGlyValGluAlaIleValAlaGluMetGlnAlaLeuGlyGInGluValPro 315
Db      143 GAGGCGACGGGGGTGAGAAATTCGTCCGACATGAGAGCCCAACGTAAGCGCGCG 202
Qy      316 ArgAspAlaPheGlyHisIleLyLeuAspAlaValAsnProGlyLyLeuTrPheGlyGlu 335
Db      203 CCGGATGCTTTGGACACGTCACGTTGGCTMAATCAATCCCGGACAGTACTCTCCCA 262
Qy      336 GlnPheAlaGInMetIleGlyValGluTrPheLeuValGInLySerGlyTyPheAla 355
Db      263 CACTTGGCGGACAAATTTGCTGCCGAAAGACATTTGCAAAAGTGGGATTCCTTGGC 322
Qy      356 ArgAlaSerAlaSerAsnValAspAspMetArgLeuIleLySerCysAlaAspLeuAla 375
Db      323 CGTCCGCGCGCGGCGCAACATTTTCATCCCACTCATGCGGCGCTGCGAGCGTGGC 382
Qy      376 ValGluCysAlaPheArgArgGluSerGlyValIleGlyHisAspGluAsnGlyAsn 395
Db      383 GTCCGCGCGGCTATTGACGGAATTCGCGATGTCAGACAGATGACAAACCAAC 442
Qy      396 Val---LeuArgAlaIleGluPheProArgIleLyGlyLyLeuProPheAsnIleAs 414
Db      443 ACCGCCATTCGAGCAATTAATTCAGTCCGATCAAGGGGAGGAAACCTTTGATATTCT 502
Qy      415 ThrAspTrpPheAsnSerMetLeuSerGluIleGlyGln 427
Db      503 CANGAATGTTCCACAAATGCTCAAGAAATTTGACAA 541

RESULT 13
CD379504      916 bp      mRNA      linear      EST 31-MAY-2003
LOCUS      PTM04673 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
DEFINITION      tricornutum cDNA 5', mRNA sequence.
ACCESSION      CD379504
VERSION      CD379504.1 GI:31255118
KEYWORDS      EST.
SOURCE      Phaeodactylum tricornutum
ORGANISM      Phaeodactylum tricornutum
Bukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
1 (bases 1 to 916)
Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.

```


THIS PAGE BLANK (uspto)